

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2005, 20:50:56 ; Search time 169 Seconds  
(without alignments)  
50.348 Million cell updates/sec

Title: US-10-776-989-9\_COPY\_20\_41  
Perfect score: 113  
Sequence: 1 BEASGVYPIDDVDYASGSGA 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 811763

Minimum DB seq length: 0  
Maximum DB seq length: 30

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : A\_Geneseq\_16Dec04:\*1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	44.2	19	3	AAY90725 Influenza
2	50	44.2	19	7	ADA08317 Scaffold
3	50	44.2	19	7	AAB38970 Insert 1
4	50	44.2	19	7	ADL10765 Structura
5	50	44.2	19	7	ADP2328 GFP test
6	50	44.2	19	8	ADK15784 Fusion pr
7	48	42.5	21	5	ABB81680 Hepatitis
8	43	38.1	22	6	ADA07146 Influenza
9	43	38.1	27	3	AAY49279 Sequence
10	41	36.3	14	2	AAY05097 HAI epito
11	41	36.3	19	6	ABP72392 Acidic ma
12	41	36.3	22	6	ABP71547 Haemagglu
13	41	36.3	27	2	AAY008426 pSMN3 fr
14	41	36.3	27	2	AAY30054 Peptide e
15	41	36.3	27	4	AAB69626 Plasma me
16	41	36.3	27	5	AAB15336 12CA5 epi
17	41	36.3	27	7	ADT100425 Peptide e
18	40	35.4	16	7	ADF70398 Orphan re
19	40	35.4	18	1	AAP20150 Influenza
20	40	35.4	21	3	AAY49278 Sequence
21	39	34.5	15	5	ABB09597 Transcrip
22	39	34.5	18	2	AAB42313 Influenza
23	39	34.5	18	2	AAR38546 Influenza
24	39	34.5	18	3	AAY87624 Influenza
25	39	34.5	18	3	AAY87628 HCV HD ep

26	39	34.5	18	3	AAB13852
27	39	34.5	18	6	ABO13659
28	39	34.5	18	8	ADR89572 Influenza
29	39	34.5	23	5	AAM47729 Peptide G
30	39	34.5	24	4	AU04463
31	39	34.5	30	7	ADC89791
32	38.5	34.1	19	6	AAC23056
33	38	33.6	16	3	AAB03963
34	38	33.6	16	3	AAB29443
35	38	33.6	16	7	ADP71037
36	38	33.6	21	3	AAY49277
37	38	33.6	21	8	ADR45760
38	38	33.6	22	4	AAM16787
39	38	33.6	22	4	ABB35771
40	38	33.6	22	4	AAM29271
41	38	33.6	22	4	ABB30607
42	38	33.6	22	4	ABB21195
43	38	33.6	22	4	AAM68960
44	38	33.6	22	4	AAM56580
45	38	33.6	22	4	ABG50623
46	38	33.6	22	4	AAM04504
47	38	33.6	22	5	ABG38546
48	38	33.6	26	8	ADR45767
49	37.5	33.2	14	2	AAY59451
50	37	32.7	8	2	AAY50222
51	37	32.7	8	2	AAY50203
52	37	32.7	11	2	AAM46950
53	37	32.7	11	2	AAM46951
54	37	32.7	12	4	AAB85134
55	37	32.7	13	3	AAB26959
56	37	32.7	13	3	AAB26958
57	37	32.7	18	2	AAM95558
58	37	32.7	20	6	ABP59398
59	37	32.7	21	4	AAB73066
60	37	32.7	23	2	AAR53019
61	37	32.7	23	2	AAR53034
62	37	32.7	25	2	AAM46332
63	37	32.7	25	2	AAM46331
64	37	32.7	27	3	AAY49280
65	37	32.7	28	2	AAR66797
66	37	32.7	30	6	ABP59412
67	37	32.7	30	6	ABP59413
68	36.5	32.3	18	8	ADL27344
69	36	31.9	8	1	AAP82985
70	36	31.9	8	2	AAY50202
71	36	31.9	9	7	ADP53331
72	36	31.9	10	2	AAR27811
73	36	31.9	10	2	AAM01229
74	36	31.9	10	2	AAM18085
75	36	31.9	10	2	AAM73266
76	36	31.9	10	2	AAM40067
77	36	31.9	10	2	AAM59362
78	36	31.9	10	2	AAM76005
79	36	31.9	10	2	AAM65242
80	36	31.9	10	2	AAM73267
81	36	31.9	10	2	AAM89255
82	36	31.9	10	3	AAY98185
83	36	31.9	10	3	AAY95076
84	36	31.9	10	4	AAB99361
85	36	31.9	10	4	AAB74593
86	36	31.9	10	4	AAB61539
87	36	31.9	10	4	AAB04320
88	36	31.9	10	4	AAB81960
89	36	31.9	10	4	AAB74994
90	36	31.9	10	4	AAB61363
91	36	31.9	10	4	AAM00829
92	36	31.9	10	4	AAB81932
93	36	31.9	10	4	AAG66444
94	36	31.9	10	4	AAG63591
95	36	31.9	10	4	AAB81962
96	36	31.9	10	5	AAG68250
97	36	31.9	10	5	AAG68245
98	36	31.9	10	5	ABG33058

AAB13852	Influenza
ABO13659	Human zin
ADR89572	Influenza
AAM47729	Peptide G
AU04463	Flag epit
ADC89791	Drosophila
AAC23056	Peptide t
AAB03963	PIAT-HA 1
AAB29443	PIAT-HA 1
ADP71037	Saccharom
AAY49277	Sequence
ADR45760	Peptide d
AAM16787	Peptide #
ABB35771	Peptide #
AAM29271	Peptide #
ABB30607	Peptide #
ABB21195	Protein #
AAM68960	Human bon
AAM56580	Human bra
ABG50623	Human liv
AAM04504	Peptide #
ABG38546	Human pep
ADR45767	Heavy cha
AAY59451	Human inf
AAY50222	Neutroph
AAY50203	Neutroph
AAM46950	Amino ter
AAM46951	Mutated a
AAB51134	Influenza
AAB26959	Mutant tr
AAB26958	Trypsino
AAM95558	Mus muscu
ABP59398	Self-asae
ABP59398	Machado-J
AAR53019	Yeast ins
AAR53034	Yeast ins
AAM46332	Haemagglu
AAM46331	Chimeric
AAY49280	Sequence
AAR66797	N-termina
ABP59412	Self-asae
ABP59413	Self-asae
ADL27344	Target ep
AAP82985	Enterokin
ADP53331	Adeno-ase
AAR27811	Tag pepti
AAM01229	Decapepti
AAM18085	Haemagglu
AAM73266	Decapepti
AAM40067	Influenza
AAM59362	Peptide 1
AAM76005	LM609 ant
AAM65242	Decapepti
AAM73267	Decapepti
AAM89255	Haemophi
AAY98185	Decapepti
AAY95076	Decapepti
AAB99361	Decapepti
AAB74593	Influenza
AAB61539	Peptide t
AAB04320	Influenza
AAB81960	Influenza
AAB74994	Decapepti
AAB61363	Peptide u
AAM00829	Haemagglu
AAB81932	Influenza
AAG66444	Influenza
AAG63591	Amino aci
AAB81962	Influenza
AAG68250	Vector de
AAG68245	Lambda ZA
ABG33058	Synthetic

99 36 31.9 10 6 ABO19801 Abol19801 7F11-alka  
100 36 31.9 10 7 ADA74778 Ada74778 Influenza

## ALIGNMENTS

RESULT 1  
ID AAY90725 standard; peptide; 19 AA.  
XX AAY90725;  
XX 16-AUG-2000 (first entry)  
XX Influenza haemagglutinin epitope tag containing 19mer peptide.  
DE Fusion; scaffold protein; peptide library; green fluorescent protein;  
KW GFP; self-binding; detection; cellular expression; screening; catabolism.  
XX Influenza virus.  
XX WO200020574-A2.  
XX 13-APR-2000.  
XX 08-OCT-1999; 99WO-US023715.  
XX 08-OCT-1998; 98US-00169015.  
XX (RIGE-) RIGEL PHARM INC.  
PI Anderson D, Bogenberger JM, Peelle BR;  
XX WPI; 2000-303765/26.  
XX Fusions of scaffold proteins with random peptide libraries for improved  
PT library screening.

XX Example 2; Page 80; 89pp; English.  
XX The present invention describes a library (X) of fusion nucleic acids,  
CC each comprising a nucleic acid encoding a random peptide, a nucleic acid  
CC encoding a scaffold protein and a nucleic acid encoding a fusion partner  
CC (the nucleic acid encoding the random peptide is inserted internally into  
CC the scaffold protein nucleic acid). Also described are: (1) a library of  
CC retroviral vectors comprising (X); (2) a library of host cells comprising  
CC (X); and (3) a method (METH) of screening for bioactive peptides  
CC conferring a particular phenotype, comprising providing cells containing  
CC (X). The constructs may be used in the screening of peptide libraries.  
CC The use of the scaffold proteins (especially green fluorescent protein  
CC (GFP)) in the fusion constructs increases cellular expression levels,  
CC increases conformational stability relative to linear peptides, decreases  
CC cellular catabolism, and increases the steady state concentrations of the  
CC random peptides and random peptide library members expressed in cells for  
CC the purposes of detecting them. The present sequence represents a peptide  
CC which comprises the influenza haemagglutinin (HA) epitope tag embedded  
CC with glycines, which is used in an example from the present invention  
XX Sequence 19 AA;

Query Match 44.2%; Score 50; DB 3; Length 19;  
Best Local Similarity 58.8%; Pred. No. 1.2;  
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 5 GYVPIDDDDYASASGSG 21  
| | | | |  
DB 3 GGYPTDVPDYASLGSGG 19

RESULT 2  
ADA08317  
ID ADA08317 standard; peptide; 19 AA.

XX ADA08317;  
XX 06-NOV-2003 (first entry)  
XX Scaffold protein fusion protein associated peptide #2.  
XX fusion nucleic acid; Renilla; GFP scaffold protein;  
KW green fluorescent protein; bioactive peptide;  
KW protein expression modification; protein folding modification;  
KW protein secretion modification; scaffold protein;  
KW cellular expression level; cellular catabolism decrease;  
KW conformational stability increase.  
XX Synthetic.  
XX US6562617-B1.  
XX 13-MAY-2003.  
XX 27-JUL-2000; 2000US-00626580.  
XX 08-OCT-1998; 98US-00169015.  
XX 08-OCT-1999; 99US-00415785.  
XX (RIGE-) RIGEL PHARM INC.  
PI Anderson D, Peelle BR, Bogenberger JM;  
XX WPI; 2003-605423/57.  
XX New library of fusion nucleic acids encoding a Renilla green fluorescent  
PT scaffold protein, a random peptide fused to the scaffold protein, and a  
PT linker, useful in e.g. modifying protein expression, protein folding or  
PT protein secretion.

XX Example 2; Col 61; 59pp; English.

XX The invention describes a library of fusion nucleic acids each  
CC comprising a first nucleic acid encoding a Renilla GFP scaffold protein;  
CC a second nucleic acid encoding a different random peptide fused to the N-  
CC terminus or C-terminus of the scaffold protein, or the second nucleic  
CC acid is inserted into the first nucleic acid; and a third nucleic acid  
CC encoding a flexible linker between the scaffold and random peptide. Also  
CC disclosed as new is a method of screening for bioactive peptides  
CC conferring a particular phenotype. The library of nucleic acids are  
CC useful in biotechnology applications, and for pharmaceutical-related  
CC applications, such as modification of protein expression, protein  
CC folding, or protein secretion. The scaffold proteins, particularly  
CC detectable genes such as GFP, are useful in fusion constructs with random  
CC and defined peptides and peptide libraries to increase the cellular  
CC expression level, decrease cellular catabolism, increase conformational  
CC stability relative to linear peptides, and to increase the steady state  
CC concentrations of the random peptides and random peptide library members  
CC expressed in cells to detect the presence of the peptides and screen  
CC random peptide libraries. This is the amino acid sequence of peptide  
CC inserted into loops structures of green fluorescent protein (GFP) and  
CC used to study the effect of structure on the fluorescence of GFP.

XX Sequence 19 AA;

Query Match 44.2%; Score 50; DB 7; Length 19;  
Best Local Similarity 58.8%; Pred. No. 1.2;  
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 5 GYVPIDDDDYASASGSG 21  
| | | | |  
DB 3 GGYPTDVPDYASLGSGG 19

RESULT 3  
AAE38970  
ID AAE38970 standard; peptide; 19 AA.

XX AAE38970;  
 AC  
 XX  
 XX  
 DT 18-DEC-2003 (first entry)  
 XX  
 DE Insert 1 peptide used in the selection of a test insert sequence.  
 XX  
 KW Bioactive peptide; cardiovascular disease; heart failure; dermatological;  
 KW arrhythmia; atherosclerosis; embolism; autoimmune disease; lupus; cancer;  
 KW neurodegenerative disease; stroke; Alzheimer's disease; psoriasis; wound;  
 KW Huntington's disease; antibiotic transport mechanism; immunosuppressive;  
 KW drug toxicity; drug resistance; osteoporosis; nootropic; anticonvulsant;  
 KW osteopathic; vulnery; anorectic; virucide; gene therapy; obesity;  
 KW cyclic.  
 XX  
 XX Unidentified.  
 OS  
 XX  
 XX  
 XX  
 PN US6548249-B1.  
 XX  
 PD 15-APR-2003.  
 XX  
 PF 27-JUL-2000; 2000US-00626581.  
 XX  
 XX 08-OCT-1998; 98US-00169015.  
 PR 08-OCT-1999; 99US-00415765.  
 XX  
 XX (RIGE-) RIGEL PHARM INC.  
 PA  
 XX  
 PI Anderson D, Peelle BR, Bogenberger JM;  
 XX  
 DR WPI; 2003-605371/57.  
 XX  
 XX Screening for bioactive peptides that may be used for treating cancer,  
 PT cardiovascular or neurodegenerative diseases, comprises providing cells  
 PT containing a fusion nucleic acid and screening the cells for an altered  
 PT phenotype.  
 XX  
 XX Example 1; Col 60; 58pp; English.  
 PS  
 XX  
 CC The invention relates to a method for screening bioactive peptides  
 CC conferring a particular phenotype. The method comprising providing cells  
 CC containing a fusion nucleic acid and screening the cells for an altered  
 CC phenotype. The method is useful in screening for bioactive peptides that  
 CC may be used to treat cancer, cardiovascular diseases (e.g. heart failure,  
 CC arrhythmia, atherosclerosis or embolism), autoimmune diseases (e.g.  
 CC lupus), neurodegenerative diseases (e.g. stroke, Alzheimer's disease or  
 CC Huntington's disease), osteoporosis, psoriasis, wounds, obesity or  
 CC infections (e.g. HIV and other viral or bacterial diseases). The method  
 CC may also be used to screen for bioactive peptides that block antibiotic  
 CC transport mechanisms, or useful in drug toxicities and drug resistance  
 CC applications. The invention is useful in gene therapy. The present  
 CC sequence is an insert peptide used in the selection of a test insert  
 CC sequence  
 XX  
 SQ Sequence 19 AA;  
 Query Match 44.2%; Score 50; DB 7; Length 19;  
 Best Local Similarity 58.8%; Pred. No. 1.2;  
 Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 5 GYPTDDDDYASASGSG 21  
 Db 3 GGYPDVDPDASLGSGG 19  
 RESULT 4  
 ADE10765  
 ID ADE10765 standard; peptide; 19 AA.  
 XX  
 AC ADE10765;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX

DE Structurally biased random peptide library related protein seqid 172.  
 XX  
 KW fusion nucleic acid library; scaffold protein; bioactive peptide;  
 KW phenotype change; cell morphology; cell growth; cell viability;  
 KW cell adhesion; cellular density; cancer; tumour; apoptosis; cell death;  
 KW loss of cell division; decreased cell growth; brca-1; brca-2;  
 KW tumour suppressor gene; breast cancer; adenomatous polyposis coli; APC;  
 KW Drosophila discs-large; Dig; cardiovascular; neurobiology; bone biology;  
 KW skin biology; cosmetic; endocrinology; infectious disease;  
 KW drug toxicity; drug resistance; inflammation; allergic response.  
 XX  
 OS Synthetic.  
 XX  
 PN US2003143562-A1.  
 XX  
 XX 31-JUL-2003.  
 XX  
 PF 20-JUN-2002; 2002US-00177725.  
 XX  
 XX 08-OCT-1998; 98US-00169015.  
 PR 08-OCT-1999; 99US-00415765.  
 XX  
 XX (RIGE-) RIGEL PHARM INC.  
 PA  
 XX  
 PI Anderson D, Peelle BR, Bogenberger JM;  
 XX  
 DR WPI; 2003-829786/77.  
 XX  
 XX Novel library of fusion nucleic acids each of which has fused first and  
 PT second nucleic acids encoding scaffold protein and library peptide having  
 PT alpha helical biasing sequence, respectively, useful in screening  
 PT methods.  
 XX  
 PS Example 2; SEQ ID NO 172; 110pp; English.  
 XX  
 CC The invention describes a library (I) of fusion nucleic acids, where each  
 CC fusion nucleic acid comprises a first nucleic acid (N1), encoding a  
 CC scaffold protein sequence; and a second nucleic acid (N2), encoding a  
 CC library peptide sequence comprising an alpha helical biasing sequence;  
 CC where N1 is fused to N2. Disclosed is a method for screening bioactive  
 CC peptides conferring a change in specific phenotype such as cell  
 CC morphology, cell growth, cell viability, adhesion to substrates or other  
 CC cells, and cellular density; changes in the expression of one or more  
 CC RNAs, proteins, lipids, hormones, cytokines, or other molecules; changes  
 CC in the equilibrium state (i.e., half-life) or one or more RNAs, protein,  
 CC lipids, hormones, cytokines, or other molecules; etc. The bioactive  
 CC peptide identified by above mentioned method is used to generate more  
 CC candidate peptides and to identify target molecules, i.e., the molecules  
 CC with which the bioactive peptide interacts. The peptide(s) can be  
 CC combined with other pharmacologic activators to study the epistatic  
 CC relationships of signal transduction pathways in question. The disclosed  
 CC method is also useful in cancer applications. Random libraries can be  
 CC introduced into any tumour cell (primary or cultured), and peptides  
 CC identified which by themselves induce apoptosis, cell death, loss of cell  
 CC division or decreased cell growth. The method is also useful for  
 CC screening of bioactive peptides which restore the constitutive function  
 CC of the brca-1 or brca-2 genes, and other tumour suppressor genes  
 CC important in breast cancer such as the adenomatous polyposis coli gene  
 CC (APC) and the Drosophila discs-large gene (Dig), which are components of  
 CC cell-cell junctions. The methods are useful in cardiovascular  
 CC applications, neurobiology applications, bone biology applications, skin  
 CC biology applications, cosmetic applications, endocrinology  
 CC applications, infectious disease applications, drug toxicities and drug  
 CC resistance applications, immunobiology, inflammation, and allergic  
 CC response applications, and biotechnology applications. The peptide  
 CC library can easily be monitored, both for its presence within cells and  
 CC its quantity. The expression of structurally biased libraries generate  
 CC elevated cellular concentration of peptides having a given structural  
 CC bias and thus increase the hit rate for targets that bind such  
 CC structures. This is the amino acid sequence of a protein associated with  
 CC fused nucleic acid and random peptide libraries of the invention.  
 XX  
 XX Sequence 19 AA;

Query Match 44.2%; Score 50; DB 7; Length 19;  
Best Local Similarity 58.8%; Pred. No. 1.2;  
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 GVPYDDDDYASASGSG 21  
| | | | |  
Db 3 GGPYDVPDYASLGSGG 19

RESULT 5  
ADF32328  
ID ADF32328 standard; peptide; 19 AA.  
AC ADF32328;  
XX  
XX  
DT 12-FEB-2004 (first entry)  
XX  
XX  
DE GFP test insert peptide.  
XX  
KW Fusion protein; green fluorescent protein; GFP; scaffold protein;  
KW cellular expression level; cellular catabolism; conformational stability;  
KW steady state concentration; arrhythmia; cardiomyocyte;  
KW contractile property; heart failure; calcium cycling; embolic phenomena;  
KW stroke; angina; myocardial infarct; atherosclerosis;  
KW low density lipoprotein; LDL; high density lipoprotein; HDL; obesity;  
KW apoptosis; neuronal function; bone morphogenic protein; hormone mimetic;  
KW fluorescence; haemagglutinin; HA protein.  
XX  
OS Synthetic.  
OS Influenza virus.  
XX  
XX  
PN US6548632-B1.  
XX  
XX  
PD 15-APR-2003.  
XX  
XX  
XX 08-OCT-1999; 99US-00415765.  
XX  
XX  
PR 08-OCT-1998; 98US-00169015.  
XX  
XX  
PA (RIG-) RIGEL PHARM INC.  
XX  
XX  
PI Anderson D, Peelle BR, Bogenberger JM;  
XX  
XX  
XX WPI; 2003-605376/57.  
XX  
XX  
PT Library of fusion proteins useful for increasing cellular expression  
PT levels, each comprises Renilla green fluorescent protein scaffold  
PT protein, random peptide fused to scaffold protein, and linker.  
XX  
XX  
PS Example 2; SEQ ID NO 57; 63pp; English.  
XX  
XX  
CC The invention discloses a library of fusion proteins (I) each comprising  
CC a Renilla green fluorescent protein (GFP) scaffold protein, a random  
CC peptide fused to the N- or C-terminus of the scaffold protein, or  
CC inserted into the scaffold protein, where each of the random peptides is  
CC different, and a flexible linker between the random peptide and the  
CC scaffold protein. (I) further comprises a second flexible linker between  
CC the other end of the random peptide and the Renilla GFP scaffold protein.  
CC The random peptide is inserted into the loop of the Renilla GFP. The  
CC linker replaces at least one amino acid of the scaffold protein. The  
CC Renilla GFP is wild-type or variant Renilla GFP. (I) is useful to  
CC increase the cellular expression levels, decrease the cellular  
CC catabolism, increase the conformational stability relative to linear  
CC peptides, and to increase the steady state concentrations of the random  
CC peptides and random peptide library members expressed in cells for the  
CC purpose of detecting the presence of the peptides and screening random  
CC peptide libraries. (I) is useful to screen a large number of cell types  
CC under a wide variety of conditions, to screen for diminished arrhythmia  
CC potential in cardiomyocytes, to screen for enhanced contractile  
CC properties of cardiomyocytes and diminished heart failure potential, to  
CC identify agents that will regulate the intracellular and sarcolemmal  
CC calcium cycling in cardiomyocytes to prevent arrhythmias, to identify

CC agents that diminish embolic phenomena in arteries and arterioles leading  
CC to strokes and angina precipitating a myocardial infarct are selected, in  
CC screening for decreases in atherosclerosis producing mechanisms to find  
CC peptides that regulate low density lipoprotein (LDL) and high density  
CC lipoprotein (HDL) metabolism, in screens to regulate obesity by the  
CC control of food intake mechanisms or diminishing the responses of  
CC receptor signaling pathways that regulate metabolism, to screen for anti-  
CC apoptotics for preservation of neuronal function and prevention of  
CC neuronal death, to screen for agonists of bone morphogenic proteins,  
CC hormone mimetics to stimulate, regulate or enhance new bone formation, in  
CC screening for bioactive peptides that block antibiotic transport  
CC mechanisms, in drug toxicities and drug resistance applications, in  
CC improving the performance of existing or developmental drugs and in  
CC biotechnology applications. The sequence presented is a test insert  
CC peptide, containing an Influenza haemagglutinin (HA) protein epitope  
CC sequence, which was used to observe the effect on GFP fluorescence.  
XX  
SQ Sequence 19 AA;

Query Match 44.2%; Score 50; DB 7; Length 19;  
Best Local Similarity 58.8%; Pred. No. 1.2;  
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 GVPYDDDDYASASGSG 21  
| | | | |  
Db 3 GGPYDVPDYASLGSGG 19

RESULT 6  
ADK15784  
ID ADK15784 standard; peptide; 19 AA.  
XX  
XX  
AC ADK15784;  
XX  
XX  
DT 06-MAY-2004 (first entry)  
XX  
XX  
DE Fusion protein library-related linker peptide #4.  
XX  
XX  
KW fusion nucleic acid library; fusion protein library; scaffold protein;  
KW green fluorescent protein; GFP; alpha helical biasing sequence;  
KW nucleating sequence; screening; linker peptide.  
XX  
XX  
OS Unidentified.  
XX  
XX  
XX US2003224412-A1.  
XX  
XX  
PD 04-DEC-2003.  
XX  
XX  
XX 18-MAR-2003; 2003US-00393449.  
XX  
XX  
PR 08-OCT-1998; 98US-00169015.  
PR 08-OCT-1999; 99US-00415765.  
PR 20-JUN-2002; 2002US-0017725.  
XX  
XX  
XX (ANDE/) ANDERSON D.  
PA (PEEL/) PEELLE B R.  
PA (BOGE/) BOGENBERGER J M.  
XX  
XX  
PI Anderson D, Peelle BR, Bogenberger JM;  
XX  
XX  
XX WPI; 2004-033956/03.  
XX  
XX  
PT Library of fusion polypeptides in which each polypeptides comprises  
PT scaffold protein and library peptide having alpha helical biasing  
PT sequence, or scaffold protein, library peptide and nucleating sequence.  
XX  
XX  
PS Example 4; SEQ ID NO 172; 110pp; English.  
XX  
XX  
CC The invention comprises a library of fusion nucleic acids, where each  
CC encoded protein contains a scaffold protein (e.g. a green fluorescent  
CC protein - GFP) and a library peptide sequence comprising an alpha helical  
CC biasing sequence, or a scaffold protein, a library peptide and a  
CC nucleating sequence. The library of the invention is useful for screening

CC bioactive peptides conferring a particular phenotype. The present amino  
 CC acid sequence represents a peptide which was used in the exemplification  
 CC of the invention.

XX Sequence 19 AA;

Query Match 44.2%; Score 50; DB 8; Length 19;

Best Local Similarity 58.8%; Pred. No. 1.2;

Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 GYVPIDDDYASAGSG 21

Db 3 GGYPDVDPYASLGSG 19

#### RESULT 7

ID ABB81680 standard; peptide; 21 AA.

XX AC ABB81680;

DT 16-AUG-2002 (first entry)

XX Hepatitis E virus ORF-2 peptide fragment HAFP.

XX Hepatitis E virus; ORF-2; HEV; virucide; immunisation; HAFP.

OS Hepatitis E virus.

XX WO200240681-A1.

XX 23-MAY-2002.

PF 30-SEP-2001; 2001WO-CN001469.

XX 30-SEP-2000; 2000CN-00130634.

XX (YANG-) YANG SHENG TANG CO LTD.

XX Xia N, Zhang J, Li S, Ge S, Gu Y, He Z;

DR WPI; 2002-427096/45.

DR N-PSDB; ABL60086.

XX Polypeptide fragments of hepatitis E virus (HEV) for vaccine compositions  
 PT and diagnosis in preventing and treating HEV infection.

XX Example 16; Page 54; 91pp; Chinese.

XX The sequence represents a fragment of hepatitis E virus (HEV) ORF-2,  
 CC shown in example 16 of the invention. The invention relates to novel n-  
 CC mer polypeptides comprising amino acids of ORF-2 of HEV where n = 2-180.  
 CC The polypeptides of the invention have virucide activity. The polypeptide  
 CC of HEV, its fragments, and chimeric proteins, are useful for vaccine  
 CC compositions and diagnosis in preventing and treating HEV infection by  
 CC immunisation, which may also be applied in the diagnosis of HEV infection

XX Sequence 21 AA;

Query Match 42.5%; Score 48; DB 5; Length 21;

Best Local Similarity 58.8%; Pred. No. 2.8;

Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4 SGVYPIDDDYASAGSG 20

Db 5 SNCYPYDVPDYASLGSG 21

#### RESULT 8

ID ADA07146 standard; peptide; 22 AA.

XX ADA07146;

AC ADA07146;

XX

DT

DE

XX

KW

KW

KW

KW

KW

XX

OS

OS

XX

PN

XX

PD

XX

PF

XX

PR

XX

PR

XX

PA

XX

PI

XX

DR

XX

XX

PT

PT

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

06-NOV-2003 (first entry)

Influenza haemagglutinin epitope tag and linker peptide.

fusion nucleic acid; Renilla; Ptilosarcus; green fluorescent protein;

GFP; retroviral vector; internal ribosome entry site element;

IRES element; type 2A sequence; protease recognition site;

antimicrobial; cytosolic; immunomodulator; antibacterial; virucide;

protozoacide; gene therapy; immunological disease; inflammatory disease;

infectious disease; cancer; drug resistance; toxicity.

Synthetic.

Influenza virus.

WO200290535-A1.

14-NOV-2002.

09-MAY-2002; 2002WO-US014766.

10-MAY-2001; 2001US-0290287P.

24-APR-2002; 2002US-00133973.

(RIGE-) RIGEL PHARM INC.

Anderson D;

WPI; 2003-103518/09.

New fusion nucleic acid, useful as reporter for intracellular assays or  
 in drug resistance or toxicity studies, comprises a gene of interest and  
 a gene encoding codon optimized Renilla muelleri or Ptilosarcus green  
 fluorescent protein.

Example 2; Page 112; 130pp; English.

The present invention describes a fusion nucleic acid (I) comprising a  
 gene of interest, and a gene encoding codon optimized Renilla muelleri or  
 Ptilosarcus green fluorescent proteins (GFP). Also described: (1) a  
 retroviral or nucleic acid vector comprising the fusion nucleic acid  
 having a promoter and a Renilla or Ptilosarcus GFP gene, and optionally,  
 a separation site, and a gene of interest; (2) a library of fusion  
 nucleic acids comprising the fusion nucleic acid described above; (3) a  
 library of cells comprising the library of fusion nucleic acids in (2);  
 (4) a cell comprising the above fusion nucleic acid; and (5) screening  
 for bioactive agents, comprising combining a candidate bioactive agent  
 and a cell comprising the fusion nucleic acid, and screening the cells  
 for an altered phenotype. Preferably the GFP is codon optimized for  
 expression in human cells. Preferably the nucleic acid vector comprises a  
 separation site comprising an internal ribosome entry site (IRES)  
 element, a type 2A sequence, or a protease recognition site. (I) has  
 antiinflammatory, cytostatic, immunomodulator, antibacterial, virucide  
 and protozoacide activities, and can be used in gene therapy. The fusion  
 nucleic acids are useful as reporters for intracellular assays, for  
 screening libraries of cells using Renilla or Ptilosarcus GFPs, in  
 immunological, inflammatory, infectious disease, or cancer applications,  
 in drug resistance or toxicity studies, or in identifying candidate  
 agents that alter phenotypes of cells. The vectors are used to express  
 the proteins and the nucleic acids in e.g. plants, insects or cells. The  
 present sequence represents a linker peptide used in an example from the  
 present invention.

Sequence 22 AA;

Query Match 38.1%; Score 43; DB 6; Length 22;

Best Local Similarity 52.9%; Pred. No. 17;

Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 GYVPIDDDYASAGSG 21

Db 4 GGYPDVDPYASLGQAG 20

RESULT 9  
AAAY49279  
ID AAY49279 standard; peptide; 27 AA.  
XX  
AC AAY49279;  
XX  
DT 07-FEB-2000 (first entry)  
XX  
XX Sequence of angiotensin HA-FLAG with 10 alanine spacer.  
XX  
XX Fusion protein; capture tag; detection tag; transgenic; gene therapy;  
KW angiotensin; FLAG peptide; HA peptide; hemagglutinin; epitope tag.  
XX  
XX Synthetic.  
XX  
PN EP960939-A2.  
XX  
XX 01-DEC-1999.  
XX  
XX 15-MAR-1999; 99EP-00105290.  
XX  
XX 23-MAR-1998; 98US-0079125P.  
XX  
XX (MEGA-) MEGABIOS CORP.  
PA (PFTZ ) PFIZER INC.  
XX  
XX Vernachio J, Papkoff J;  
PI  
XX WPI; 2000-015438/02.  
DR  
XX  
XX New fusion protein, useful for detection of protein in transgenic animal.  
PT  
XX  
PS Disclosure; Fig 1; 17pp; English.  
XX  
XX The invention relates to a fusion protein comprising a polypeptide  
CC sequence of interest, a capture tag sequence and a detection tag  
CC sequence, where both the capture tag sequence and the detection tag  
CC sequence are homologous to the polypeptide of interest. The nucleic  
CC acid encoding the fusion protein may be used to generate transgenic  
CC animals expressing the fusion protein. The protein may then be detected  
CC in these animals by screening samples from the animals with antibodies  
CC that bind to the capture tag and detection tag sequences. This method may  
CC be used to detect gene products expressed as a result of gene therapy,  
CC especially when used to express angiotensin in mice. The nucleic acid may  
CC also be ligated to expression vectors to allow transformation of cell  
CC cultures. The structure of the new fusion protein allows detection of the  
CC protein of interest when expressed in a transgenic animal, more easily  
CC than prior art fusion proteins. The new method of detection is fast,  
CC cheap and easy to carry out compared to prior art methods. Sequences  
CC AAY49277-80 represent C-terminus double tagged angiotensin proteins of  
CC the invention, where the capture tag sequence is a FLAG peptide and the  
CC detection tag sequence is a HA peptide derived from influenza  
CC hemagglutinin protein  
XX  
SQ Sequence 27 AA;  
Query Match 38.1%; Score 43; DB 3; Length 27;  
Best Local Similarity 50.0%; Pred. No. 22;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
QY 7 YPIDDDDYASGSGA 22  
DB 1 YPVDVDPYAAAAA 16  
RESULT 10  
AAW05097  
ID AAW05097 standard; peptide; 14 AA.  
XX  
AC AAW05097;  
XX  
XX 10-JUN-1997 (first entry)  
DT

XX  
DE HA1 epitope.  
XX  
KW HA1 epitope; influenza virus; haemagglutinin protein; vector; human;  
KW pX1-CMV-Gax-HA; early promoter; enhancer; cytomegalovirus; CMV; rabbit;  
KW leader sequence; herpes simplex virus; thymidine kinase; rat; Ad5; Ad2;  
KW gax; polyadenylation signal; beta-globin gene; E4 region; gene therapy;  
KW growth arrest-specific homeobox protein; adenovirus; E1 region;  
KW replication; vascular hyperproliferative disorder;  
KW post-angioplastic restenosis.  
XX  
OS Influenza virus.  
XX  
PN WO9630385-A1.  
XX  
PD 03-OCT-1996.  
XX  
PF 28-MAR-1996; 96WO-US004493.  
XX  
PR 31-MAR-1995; 95PR-00004234.  
XX  
XX (UYCA-) UNIV CASE WESTERN RESERVE.  
PA  
PI Branellec D, Walsh K, Isner JM, Denefle P;  
XX WPI; 1996-455263/45.  
XX  
PT Recombination-defective viral vector contg. gax gene - for gene therapy  
PT of hyperproliferative disorders, esp. restenosis.  
XX  
PS Example 1; Page 22; 59pp; English.  
XX  
CC This sequence represents the HA1 epitope from the influenza virus  
CC haemagglutinin protein. This epitope was included in the construction of  
CC the vector pX1-CMV-Gax-HA. This vector also comprises the early promoter  
CC and enhancer sequence of cytomegalovirus, the leader sequence of herpes  
CC simplex virus thymidine kinase, including the AUG codon and first 3 amino  
CC acids, the rat gax cDNA and the polyadenylation signal of the rabbit beta  
CC-globin gene. The resulting vector was used in the construction of a  
CC defective recombinant virus contg. at least one inserted gene encoding  
CC all or part of a GAX (growth arrest-specific homeobox) protein. Pref. the  
CC virus is an adenovirus, pref. of Ad5 or Ad2 type, lacking genomic regions  
CC necessary for its replication in infected cells, esp. all or part of the  
CC E1 region and opt. the E4 region. The adenovirus is of animal (esp.  
CC canine) origin. The virus could also be an adeno-associated virus or a  
CC retrovirus. The inserted gene encodes the rat GAX protein or its human  
CC homologue. The inserted gene is cDNA or genomic DNA and is linked to  
CC sequences enabling it to be expressed in infected cells. The defective  
CC virus is useful for gene therapy of vascular hyperproliferative  
CC disorders, esp. post-angioplastic restenosis, pref. in injectable doses  
CC of 104-1014 pfu/ml  
XX  
SQ Sequence 14 AA;  
Query Match 36.3%; Score 41; DB 2; Length 14;  
Best Local Similarity 61.5%; Pred. NO. 21;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 7 YPIDDDDYASASG 19  
DB 1 YPVDVDPYASLGG 13  
RESULT 11  
ABP72392  
ID ABP72392 standard; peptide; 19 AA.  
XX  
AC ABP72392;  
XX  
XX 13-MAY-2003 (first entry)  
XX  
XX Acidic mammalian chitinase peptide.  
XX

KW Acidic mammalian chitinase; enzyme; mouse; transgenic animal; inhibitor;  
 KW antiinflammatory; gastrointestinal; dermatological; antiallergic;  
 KW antiasthmatic.  
 OS Mus sp.  
 XX WO2003009808-A2.  
 XX 06-FEB-2003.  
 XX 23-JUL-2002; 2002WO-US023516.  
 XX 24-JUL-2001; 2001US-0307432P.  
 XX (UYVA ) UNIV YALE.  
 PA (ELIA/) ELIAS J A.  
 PA (ZHUZ/) ZHU Z.  
 XX Elias JA, Zhu Z;  
 XX WPI; 2003-289788/28.  
 XX Treating an inflammatory disease associated with increased levels of  
 PT chitinase and chitinase-like molecules, such as chronic obstructive  
 PT pulmonary disease, pneumonia, inflammatory bowel disease, emphysema and  
 PT atopic dermatitis.  
 XX Example; Page 47; 100pp; English.  
 XX The present sequence is that of a peptide from acidic mammalian chitinase  
 CC AMCase. The peptide was used to immunise rabbits in order to generate  
 CC polyclonal antibodies to AMCase. The antibodies inhibited antigen-induced  
 CC inflammation in the lungs of OVA-sensitized mice (mouse model of human  
 CC asthma). The invention provides methods for treating or preventing an  
 CC inflammatory disease in a mammal (human), where the disease is associated  
 CC with an increased level of a chitinase-like molecule (e.g. AMCase) or  
 CC interleukin-13, or with a Th2 inflammatory response. The methods involve  
 CC the administration of an inhibitor of a chitinase-like molecule, such as  
 CC an antibody, ribozyme or antisense nucleic acid. The inflammatory disease  
 CC is chronic obstructive pulmonary disease (COPD), interstitial lung  
 CC disease, chronic obstructive lung disease, chronic bronchitis,  
 CC eosinophilic bronchitis, eosinophilic pneumonia, pneumonia, inflammatory  
 CC bowel disease, atopic dermatitis, atopy, allergy, allergic rhinitis,  
 CC idiopathic pulmonary fibrosis, scleroderma, emphysema or asthma (all  
 CC claimed). Methods are also provided for identifying a compound useful for  
 CC treating an inflammatory disease  
 XX Sequence 19 AA;  
 SQ  
 Query Match 36.3%; Score 41; DB 6; Length 19;  
 Best Local Similarity 50.0%; Pred. No. 30;  
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 ERASGVPTDDDD 12  
 Db ::|||::||  
 2 DRADGLYPVADD 13  
 RESULT 12  
 ABP71547  
 ID ABP71547 standard; peptide; 22 AA.  
 XX AC  
 XX ABP71547;  
 XX 29-MAY-2003 (first entry)  
 DE Haemagglutinin epitope (HA) peptide.  
 XX Phosphoinositide 3-kinase; PI3K; phosphatidylinositol-3,4,5-P-3; PI3P;  
 KW GTPase; GTPase-responsive domain; GRD; p85; cytoskeletal; cancer; HA;  
 KW haemagglutinin.  
 XX Unidentified.

XX WO2003006487-A2.  
 XX 23-JAN-2003.  
 XX 11-JUL-2002; 2002WO-US022153.  
 XX 11-JUL-2001; 2001US-0304498P.  
 PR 30-JUL-2001; 2001US-0308654P.  
 PR 11-MAR-2002; 2002US-0363078P.  
 XX (UYJE-) UNIV JEFFERSON THOMAS.  
 XX Chan TO;  
 PI WPI; 2003-239233/23.  
 DR Modulation of phosphoinositide 3-kinase activity or phosphatidylinositol-  
 XX 3,4,5-P3 production useful for treating cancer involves abrogating GTPase  
 PT -responsive domain in p85 subunit.  
 XX Example 7; Page 28; 53pp; English.  
 XX The invention relates to the modulation of phosphoinositide 3-kinase  
 CC (PI3K) activity or phosphatidylinositol-3,4,5-P-3 (PI3P) production. The  
 CC method involves abrogating GTPase-responsive domain (GRD) in p85 subunit  
 CC of (PI3K) activity or overexpressing only inhibitory domain portion of  
 CC p85 subunit of (PI3K). The method is useful for modulating PI3K activity  
 CC or PI3P production; for modulating ras-induced PI3P synthesis in a  
 CC mammalian cell (e.g. human); for selecting modulator of PI3P induction by  
 CC a small GTPase; for treating cancer. The present sequence represents a  
 CC haemagglutinin epitope (HA) peptide  
 XX Sequence 22 AA;  
 SQ  
 Query Match 36.3%; Score 41; DB 6; Length 22;  
 Best Local Similarity 61.5%; Pred. No. 36;  
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 7 YPIDDDDVASAG 19  
 Db |||||  
 5 YPYDVPDYASLGG 17  
 RESULT 13  
 AAW08426  
 ID AAW08426 standard; peptide; 27 AA.  
 XX AC AAW08426;  
 XX 03-OCT-1997 (first entry)  
 DE pSMTN3 fragment for cloning transcription factor fusion proteins.  
 XX Rapamycin; responsive transfectant cell; recombinant DNA; fusion protein;  
 KW rapamycin binding domain; RBD; FKBP; FK506-binding protein; FRB domain;  
 KW genetic engineering; regulation; transcription; signal transduction;  
 KW apoptosis; inhibition; virus; gene therapy.  
 XX Synthetic.  
 OS Key Location/Qualifiers  
 FH Region 4..19  
 FT /product= "12CA5\_epitope"  
 FT Region 21..27  
 FT /product= "SV40\_T\_NLS"  
 XX WO9641865-A1.  
 XX 27-DEC-1996.  
 XX 07-JUN-1996; 96WO-US009948.  
 XX



PR 07-JUN-1995; 95US-00481941.  
 PR 09-FEB-1996; 96US-00598776.  
 XX  
 PA (ARIA-) ARIAD GENE THERAPEUTICS INC.  
 XX  
 PI Clackson T, Holt DA, Gilman MZ;  
 XX  
 DR WPI; 1997-065453/06.  
 DR N-PSDB; AAT49029.  
 XX  
 XX New rapamycin-responsive transfected cells - contg. recombinant DNA  
 PT encoding fusion proteins which act as biological switches for regulating  
 PT biological events.  
 XX  
 PS Example 1; Page 65; 149pp; English.  
 XX  
 CC The present sequence represents the encoding DNA sequence inserted into  
 CC pSMTN2 to produce pSMTN3, which is used for the cloning of transcription  
 CC factor fusion proteins. This is used in the production of animal cells  
 CC containing at least two recombinant DNAs. One DNA should encode a  
 CC chimeric protein which is capable of binding to rapamycin, or a rapamycin  
 CC analogue, and comprises at least one rapamycin binding domain (RBD) and  
 CC at least one protein domain heterologous to it; the second recombinant  
 CC DNA should encode a second chimeric protein which is capable of forming a  
 CC complex with rapamycin, or a rapamycin analogue, and the first chimeric  
 CC protein and comprises at least one FKBP:rapamycin binding (FRB) domain  
 CC and at least one domain heterologous to it. Contacting these genetically  
 CC engineered cells with rapamycin or analogues results in the formation of  
 CC a complex between the fusion proteins, and the rapamycin), and  
 CC initiation of a biological response. The products can be used for  
 CC regulating biological events such as gene transcription, activation of an  
 CC intracellular signal transduction pathway leading to e.g. gene expression  
 CC or apoptotic cell death, gene knock-out, blockade of a gene or inhibition  
 CC of function of a gene product. They are used particularly for regulated  
 CC gene therapy and for production of recombinant proteins and viruses  
 XX  
 SQ Sequence 27 AA;  
 Query Match 36.3%; Score 41; DB 2; Length 27;  
 Best Local Similarity 61.5%; Pred. No. 45;  
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 7 YPIDDDDYASAG 19  
 Db ||| |||||  
 5 YPYDVPDYASLGG 17  
 RESULT 14  
 AAY30054  
 ID AAY30054 standard; peptide; 27 AA.  
 XX  
 AC AAY30054;  
 XX  
 DT 06-OCT-1999 (first entry)  
 XX  
 DE Peptide encoded by a DNA sequence used to create plasmid pSMTN3.  
 XX  
 KW Multimerisation; chimeric protein; rapamycin compound; gene therapy;  
 KW transcription; cell growth; cell proliferation; cell differentiation;  
 KW apoptosis.  
 XX  
 OS Synthetic.  
 OS  
 PN WO9936553-A2.  
 XX  
 PD 22-JUL-1999.  
 XX  
 PF 15-JAN-1999; 99WO-US000178.  
 XX  
 PR 15-JAN-1998; 98US-0071591P.  
 PR 21-JAN-1998; 98US-0072016P.  
 PR 22-JAN-1998; 98US-00012097.  
 PR 22-JAN-1998; 98US-0072219P.

XX  
 PA (ARIA-) ARIAD GENE THERAPEUTICS INC.  
 XX  
 PI Clackson TP, Gilman MZ, Holt DA, Keenan TP, Rozamus L, Yang W;  
 XX  
 DR WPI; 1999-468986/39.  
 DR N-PSDB; AAX86602.  
 XX  
 XX Multimerizing chimeric proteins in cells useful in gene therapy, for e.g.  
 PT target gene transcription and growth of engineered cells.  
 XX  
 PS Example 7; Page 108; 155pp; English.  
 XX  
 CC The specification describes a process for multimerising chimeric proteins  
 CC in cells. The method comprises forming a complex between a rapamycin  
 CC compound and cells containing at least one recombinant nucleic acid (rNA)  
 CC construct encoding specialised chimeric proteins (Cps). The method is  
 CC used for multimerising chimeric proteins in cells, which is useful for  
 CC gene therapy, activating the transcription of a desired gene, actuating  
 CC cell growth, proliferation, differentiation or apoptosis, or triggering  
 CC other biological events in engineered cells in a rapalog-dependent  
 CC manner. The present sequence is used to create plasmid pSMTN3, which is  
 CC used to express the chimeric genes of the invention  
 XX  
 SQ Sequence 27 AA;  
 Query Match 36.3%; Score 41; DB 2; Length 27;  
 Best Local Similarity 61.5%; Pred. No. 45;  
 Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 7 YPIDDDDYASAG 19  
 Db ||| |||||  
 5 YPYDVPDYASLGG 17  
 RESULT 15  
 AAB69626  
 ID AAB69626 standard; protein; 27 AA.  
 XX  
 AC AAB69626;  
 XX  
 DT 30-APR-2001 (first entry)  
 XX  
 DE Plasma membrane targeting signal peptide.  
 XX  
 KW Neurological disorder; Huntington's disease; Alzheimer's disease;  
 KW Parkinson's disease; prion disease; frontotemporal dementia;  
 KW amyotrophic lateral sclerosis; spinal and bulbar muscular atrophy;  
 KW dentatorubal-pallidoluyisan atrophy; spinocerebellar ataxia type 1; SCA2;  
 KW SCA3; SCA4; SCA5; SCA6; SCA7; protein accumulation; intrabody.  
 XX  
 OS Synthetic.  
 OS  
 PN WO200106989-A2.  
 XX  
 PD 01-FEB-2001.  
 XX  
 PF 24-JUL-2000; 2000WO-US020131.  
 XX  
 PR 27-JUL-1999; 99US-0146047P.  
 PR 21-JUL-2000; 2000US-00620955.  
 XX  
 PA (HUST/) HUSTON J S.  
 PA (MESS/) MESSER A.  
 PA (LECE/) LECERF J.  
 XX  
 PI Huston JS, Messer A, Lecerf J;  
 XX  
 DR WPI; 2001-182700/18.  
 XX  
 PT Inhibiting intracellular polypeptide accumulation, useful for treating  
 PT neurological disorders, e.g. Alzheimer's disease, comprises contacting  
 PT the polypeptide with a specific intrabody.



XX PS Disclosure; Fig 2; 108pp; English.

XX CC The present invention describes a method for inhibiting the formation of

CC aggregates of certain proteins, involving contacting the protein with a

CC binding molecule known as an intrabody. Proteins to be bound include

CC those associated with neurological disorders, and so the method can be

CC used in the prevention of diseases such as Alzheimer's, Parkinson's and

CC Huntington's diseases, prion disease, frontotemporal dementia,

CC amyotrophic lateral sclerosis, spinal and bulbar muscular atrophy,

CC dentatorubral-pallidoluysian atrophy, spinocerebellar ataxia type 1

CC (SCA1), SCA2, SCA3, SCA4, SCA5, SCA6 and SCA7

XX SQ Sequence 27 AA;

Query Match 36.3%; Score 41; DB 4; Length 27;

Best Local Similarity 61.5%; Pred. No. 45;

Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 7 YPIDDDDYASAG 19

Db ||| |||||

1 YPVDVDPYASKDG 13

RESULT 16

ID AAE15336

XX AAE15336 standard; protein; 27 AA.

AC AAE15336;

XX 07-MAR-2002 (first entry)

DE 12CA5 epitope fragment #4.

XX Chimeric transcription factor; nuclear factor kappaB; NF-kappaB p65;

KW biological material; biological research; biochemical purification;

KW transgenic animal; gene therapy; acidic activation motif.

XX Unidentified.

OS US6306649-B1.

PN 23-OCT-2001.

XX 27-JUN-1996; 96US-00672213.

XX 27-JUN-1995; 95US-0000553P.

PR 29-DEC-1995; 95US-0019614P.

XX (ARIA-) ARIAD GENE THERAPEUTICS INC.

PA Gilman MZ, Natesan S;

PI WPI; 2002-033255/04.

XX N-PSDB; AAD24545.

DR New recombinant DNA sequence encoding a chimeric transcription factor,

PT useful in gene therapy by elevating the expression of therapeutic target

PT gene and allowing expression to reach therapeutically levels.

XX Example 2; Col 31-32; 47pp; English.

XX The invention relates to a recombinant DNA sequence encoding a chimeric

CC transcription factor. The transcription factor comprises one or more

CC copies of a peptide sequence comprising all or part of a peptide sequence

CC spanning 361-550 of human nuclear factor (NF)-kappaB p65 and a peptide

CC sequence heterologous to the sequence selected from VP16 V8, VP16 B, VP16

CC C, HSF or CTF. The recombinant DNA is used for producing biological

CC materials, in biological researches where precise control over a target

CC gene is desired, such as in the expression of a protein or RNA of

CC interest for biochemical purification and in tissue or organ specific

CC expression of a protein or RNA in transgenic animals to elevate its

CC biological function. It is also useful in gene therapy strategies by

XX CC substantially elevating the expression of therapeutic target gene and

CC allowing expression to reach therapeutically effective levels. The

CC transcription factors are useful for effecting transcription of target

CC genes in genetically engineered cells or organisms containing them. The

CC present sequence is 12CA5 epitope fragment used in the exemplification of

CC the invention

XX SQ Sequence 27 AA;

Query Match 36.3%; Score 41; DB 5; Length 27;

Best Local Similarity 61.5%; Pred. No. 45;

Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 7 YPIDDDDYASAG 19

Db ||| |||||

5 YPVDVDPYASLGG 17

RESULT 17

ID ADI00425

XX ADI00425 standard; peptide; 27 AA.

AC ADI00425;

XX 22-APR-2004 (first entry)

DE Peptide encoded by DNA inserted into pSMTN2 to create pSMTN3 vector.

KW gene expression; transcription factor; gene therapy; pSMTN2; pSMTN3;

KW SV40.

XX Simian virus 40.

OS Synthetic.

OS Unidentified.

XX US2003143731-A1.

PN 31-JUL-2003.

XX 23-OCT-2001; 2001US-00002244.

XX 27-JUN-1995; 95US-0000553P.

PR 29-DEC-1995; 95US-0019614P.

XX 27-JUN-1996; 96US-00672213.

PR 11-JUN-1998; 98US-00096732.

PR 04-MAR-1999; 99US-00262721.

PR 28-SEP-1999; 99US-00407402.

XX (ARIA-) ARIAD GENE THERAPEUTICS INC.

PA Gilman MZ, Natesan S;

PI WPI; 2003-678757/64.

DR N-PSDB; ADI00426.

DR Expressing a target gene in a cell within a host organism, useful in gene

PT therapy, comprises introducing into the organism (cells containing) a

PT transcription factor construct and a target gene construct.

XX Example; SEQ ID NO 33; 35pp; English.

XX The invention relates to a novel method for expressing a target gene in a

CC cell within a host organism comprising introducing cells containing a

CC first DNA construct encoding a transcription factor and a second

CC construct encoding a target gene operably linked to a promoter and

CC transcription control sequences recognised by the transcription factor of

CC the first construct. The methods and constructs of the invention may be

CC useful for the expression of target genes in genetically engineered cells

CC or organisms containing them, as well as in gene therapy, for producing

CC biological materials or in biological research. The current sequence is

CC that of the peptide encoded by the DNA fragment inserted into the pSMTN2

CC vector to create the pSMTN3 vector DNA of the invention.

SQ Sequence 27 AA;

Query Match 36.3%; Score 41; DB 7; Length 27;  
Best Local Similarity 61.5%; Pred. No. 45;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 YPIDDDDDYASAG 19  
| | | | |  
DB 5 YPYDVPDYASLGG 17

RESULT 18

ADF70398  
ID ADF70398 standard; peptide; 16 AA.

XX AC

XX AC

XX DT

XX 12-FEB-2004 (first entry)

DE Orphan receptor ligand-related HA tag.

KW ligand; orphan receptor protein; fusion protein; fluorescent protein;  
cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;  
GFPuv; Enhanced GFP; EGFP; HA tag.

OS Unidentified.

XX WO2003071272-A1.

XX 28-AUG-2003.

XX 21-FEB-2003; 2003WO-JF001901.

XX 22-FEB-2002; 2002JP-00045728.

XX 23-JUL-2002; 2002JP-00213949.

XX 11-OCT-2002; 2002JP-00298237.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;  
WPI; 2003-697654/66.

PT Transformation of cells with a fusion protein of an orphan receptor  
protein with a fluorescent protein useful for identification of ligands  
to the orphan receptor.

PS Disclosure; SEQ ID NO 21; 594pp; Japanese.

CC This invention relates to a novel method of identifying ligands to an  
orphan receptor protein which comprises transforming cells with DNA  
encoding a fusion protein of the orphan receptor with a fluorescent  
protein, so that the fusion protein is expressed in the cells (or cell  
membranes isolated from them) and contacting the cells with the potential  
ligand to be tested. A suitable fluorescent protein for incorporation in  
the fusion protein is green fluorescent protein (GFP), for example GFP-1,  
wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the  
identification of ligands binding to an orphan receptor protein.

SQ Sequence 16 AA;

Query Match 35.4%; Score 40; DB 7; Length 16;  
Best Local Similarity 66.7%; Pred. No. 35;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 GYVPIDDDDDYAS 16

DB 2 GSYPYDVPDYAS 13

RESULT 19

AAP20150

ID AAP20150 standard; protein; 18 AA.

XX AAP20150;  
AC 06-JUL-1992 (first entry)  
DT Influenza virus antigen.  
DE Influenza virus antigen.  
XX Influenza virus; antigen; vaccine.  
XX Synthetic.  
XX DE3200813-A.  
XX 12-AUG-1982.

XX 13-JAN-1981; 81IL-00061904.

XX 13-JAN-1981; 81IL-00061904.

XX (YEDA ) YEDA RES & DEV CO LTD.

XX Arnon R, Shapira M, Mueller G;  
WPI; 1982-68417E/33.

XX synthetic vaccine for virus infections - contg. synthetic peptide virus  
antigen fragment on carrier, esp. synthetic influenza antigen on tetanus  
toxoid carrier.  
XX Claim 5; Page 2; 16pp; German.  
XX The synthetic peptide corresponds to an antigen fraction of influenza  
virus, and is attached to a carrier for use as a vaccine against  
influenza

SQ Sequence 18 AA;

Query Match 35.4%; Score 40; DB 1; Length 18;

Best Local Similarity 61.5%; Pred. No. 40;

Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 SGVPIDDDDYAS 16

DB 5 SNAYPYDVPDYAS 17

RESULT 20

AAY49278  
ID AAY49278 standard; peptide; 21 AA.

XX AC

XX AAY49278;

XX 07-FEB-2000 (first entry)

XX Sequence of angiotensin FLAG-HA with 4 alanine spacer.

XX Fusion protein; capture tag; detection tag; transgenic; gene therapy;  
angiotensin; FLAG peptide; HA peptide; hemagglutinin; epitope tag.  
XX Synthetic.  
XX EP960939-A2.  
XX 01-DEC-1999.  
XX 15-MAR-1999; 99EP-00105290.  
XX 23-MAR-1998; 98US-0079125P.  
XX (MEGA-) MEGABIOS CORP.  
XX (PFIZ ) PFIZER INC.  
XX Vernachio J, Papkoff J;

XX WPI; 2000-015438/02.  
 XX New fusion protein, useful for detection of protein in transgenic animal.  
 PT Disclosure; Fig 1; 17pp; English.  
 XX The invention relates to a fusion protein comprising a polypeptide  
 CC sequence of interest, a capture tag sequence and a detection tag  
 CC sequence, where both the capture tag sequence and the detection tag  
 CC sequence are heterologous to the polypeptide of interest. The nucleic  
 CC acid encoding the fusion protein may be used to generate transgenic  
 CC animals expressing the fusion protein. The protein may then be detected  
 CC in these animals by screening samples from the animals with antibodies  
 CC that bind to the capture tag and detection tag sequences. This method may  
 CC be used to detect gene products expressed as a result of gene therapy,  
 CC especially when used to express angiotensin in mice. The nucleic acid may  
 CC also be ligated to expression vectors to allow transformation of cell  
 CC cultures. The structure of the new fusion protein allows detection of the  
 CC protein of interest when expressed in a transgenic animal, more easily  
 CC than prior art fusion proteins. The new method of detection is fast,  
 CC cheap and easy to carry out compared to prior art methods. Sequences  
 CC AAY49277-80 represent C-terminus double tagged angiotensin proteins of  
 CC the invention, where the capture tag sequence is a FLAG peptide and the  
 CC detection tag sequence is a HA peptide derived from influenza  
 CC hemagglutinin protein  
 XX Sequence 21 AA;  
 SQ Query Match 35.4%; Score 40; DB 3; Length 21;  
 Best Local Similarity 46.7%; Pred. No. 48;  
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 EASGVYPIDDDDYA 15  
 DB 7 DKAAAYPIVDVDDYA 21  
 RESULT 21  
 ID ABB09597  
 XX ABB09597 standard; peptide; 15 AA.  
 AC ABB09597;  
 XX 24-SEP-2002 (first entry)  
 DT Transcription regulation protein 34.54 N-terminal peptide.  
 DE Transcription regulation protein 34.54; human; recombinant production;  
 KW malignant tumour; cancer; blood disease; HIV infection; gene therapy;  
 KW human immunodeficiency virus; immune disorder; inflammatory condition;  
 KW cytostatic; antiinflammatory; immunomodulator; N-terminal peptide;  
 KW enzyme linked immunosorbent assay; ELISA.  
 XX Homo sapiens.  
 OS CN1339448-A.  
 PN 13-MAR-2002.  
 PD 21-AUG-2000; 2000CN-00119658.  
 PF 21-AUG-2000; 2000CN-00119658.  
 XX 21-AUG-2000; 2000CN-00119658.  
 XX (BODE-) BODE GENE DEV CO LTD SHANGHAI.  
 PA Mao Y, Xie Y;  
 PI WPI; 2002-455350/49.  
 DR New polypeptide-transcription regulation protein 34.54 and polynucleotide  
 PT encoding the polypeptide.  
 XX

PS Example 6; Page 19 (Disclosure); 34pp; Chinese.  
 XX The invention relates to transcription regulation protein 34.54  
 CC (ABB09596) and nucleic acids encoding it (ABL55498). The protein has a  
 CC molecular weight of 34.54 kD and has 51% identity and 73% homology over a  
 CC 308 amino acid stretch with the protein given in ABB09598. The invention  
 CC also relates to a method for the recombinant production of the protein,  
 CC an antagonist of the protein, and the use of the protein, gene and  
 CC antagonist in therapeutic applications. Transcription regulation protein  
 CC 34.54 can be used in the treatment of a variety of diseases such as  
 CC malignant tumours, blood diseases, HIV (human immunodeficiency virus)  
 CC infection, immune disorders and inflammatory conditions. The present  
 CC sequence represents the 15 N-terminal amino acids of transcription  
 CC regulation protein 34.54 used in ELISA (enzyme linked immunosorbent  
 CC assay) in an exemplification of the invention  
 XX Sequence 15 AA;  
 SQ Query Match 34.5%; Score 39; DB 5; Length 15;  
 Best Local Similarity 45.5%; Pred. No. 47;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 SGVYPIDDDDY 14  
 DB 4 AGMYPIDQSEY 14  
 RESULT 22  
 ID AAR42313  
 XX AAR42313 standard; protein; 18 AA.  
 AC AAR42313;  
 XX 25-MAR-2003 (revised)  
 DT 17-MAY-1994 (first entry)  
 DE Influenza virus HA epitope 91-108.  
 KW Haemagglutinin; chimera; nucleoprotein; vaccine; mucosal; systemic;  
 KW immunisation; adjuvant free.  
 XX Influenza virus.  
 OS WO9320846-A1.  
 PN 28-OCT-1993.  
 PD 19-APR-1993; 93WO-US003606.  
 PF 17-APR-1992; 92IL-00101639.  
 XX (ARNO/) ARNON Y.  
 PA (YEDA ) YEDA RES & DEV CO LTD.  
 XX Arnon R, Levi R;  
 PI WPI; 1993-351370/44.  
 DR N-PSDB; AAQ51012.  
 XX Synthetic recombinant influenza vaccines - comprises chimeric proteins,  
 PT having flagellin and haemagglutinin or nucleoprotein components.  
 PS Claim 3; Page 24; 45pp; English.  
 XX The sequence is that of an epitope of influenza virus haemagglutinin from  
 CC residues 91-108. The sequence can be used in a chimeric protein also  
 CC comprising the flagellin amino acid sequence. Such a protein can be used  
 CC in a vaccine compen. for mucosal and systemic immunisation against  
 CC influenza virus infections. The vaccines do not require adjuvants and may  
 CC therefore be administered intranasally. See also AAR42314-5. (Updated on  
 CC 25-MAR-2003 to correct PN field.)  
 XX Sequence 18 AA;  
 SQ

DT	21-JUL-2000	(first entry)
XX		
DE	Influenza virus hydrophobic domain epitope #3.	
XX		
KW	Immunogenic liposome; antigen; cytostatic; hydrophobic domain; epitope;	
XX	vesicle forming lipid; immunization; vaccine; antimicrobial.	
XX		
OS	Influenza virus.	
XX		
FN	WO200016746-A2.	
XX		
PD	30-MAR-2000.	
XX		
PF	22-SEP-1999; 99WO-US020880.	
XX		
PR	22-SEP-1999; 98US-0101351P.	
XX		
PR	21-SEP-1999; 99US-0400723P.	
XX		
PA	(MOLE-) MOLECULAR EXPRESS INC.	
XX		
PI	Fujii G, Cramer DV, Ernst WA, Adler-Moore J, Perry LJ;	
XX		
DR	WPI; 2000-283420/24.	
XX		
PT	New immunogenic liposome composition for inducing an immunogenic response	
PT	in a host animal against a variety of microbial agents and cancers,	
PT	comprises vesicle forming lipids and an antigenic construct.	
XX		
PS	Example IV; Page 26; 44pp; English.	
XX		
CC	This invention describes a novel immunogenic liposome composition (C1)	
CC	comprising vesicle forming lipids and an antigenic construct comprising	
CC	one or more antigenic determinants and a hydrophobic domain (HD)	
CC	associated with the membrane of the liposome composition. C1 is used to	
CC	induce an immunogenic response in a host animal, preferably birds or	
CC	mammals such as humans, against a variety of microbial agents and	
CC	cancers. C1 provides effective and safe immunization. The epitopes can be	
CC	readily changed to provide for maximal flexibility in vaccine design.	
CC	Also, the expressed carrier protein is water soluble and can be easily	
CC	purified using standard protein preparation methods. The products of the	
CC	invention have antimicrobial and cytostatic activity. This sequence	
CC	represents an influenza virus HD epitope which is used to illustrate the	
CC	method of the invention	
XX		
SQ	Sequence 18 AA;	
	Query Match	34.5%; Score 39; DB 3; Length 18;
	Best Local Similarity	61.5%; Pred. No. 58;
	Matches	8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy	4 SGVYPIDDDDYAS 16	
Db	5 SNCYPIVDVDYAS 17	
RESULT 25		
AAV87628		
ID	AAV87628 standard; protein; 18 AA.	
XX		
AC	AAV87628;	
XX		
DT	21-JUL-2000 (first entry)	
DE	HCV HD epitope #3.	
XX		
KW	Cytokine; immunogenic liposome; antigen; cytostatic; hydrophobic domain;	
XX	vesicle forming lipid; immunization; vaccine; antimicrobial.	
XX		
OS	Hepatitis C virus.	
XX		
FN	WO200016746-A2.	
XX		
PD	30-MAR-2000.	

XX 22-SEP-1999; 99WO-US020880.  
XX  
PR 22-SEP-1998; 98US-0101351P.  
PR 21-SEP-1999; 99US-0400723P.  
XX  
PA (MOLE-) MOLECULAR EXPRESS INC.  
XX  
PI Fujii G, Cramer DV, Ernst WA, Adler-Moore J, Perry LJ;  
XX WPI; 2000-283420/24.  
DR  
XX  
PT New immunogenic liposome composition for inducing an immunogenic response  
PT in a host animal against a variety of microbial agents and cancers,  
PT comprises vesicle forming lipids and an antigenic construct.  
XX  
PS Example VI; Page 34; 4pp; English.  
XX  
CC This invention describes a novel immunogenic liposome composition (C1)  
CC comprising vesicle forming lipids and an antigenic construct comprising  
CC one or more antigenic determinants and a hydrophobic domain (HD)  
CC associated with the membrane of the liposome composition. C1 is used to  
CC induce an immunogenic response in a host animal, preferably birds or  
CC mammals such as humans, against a variety of microbial agents and  
CC cancers. C1 provides effective and safe immunization. The epitopes can be  
CC readily changed to provide for maximal flexibility in vaccine design.  
CC Also, the expressed carrier protein is water soluble and can be easily  
CC purified using standard protein preparation methods. The products of the  
CC invention have antimicrobial and cytostatic activity. This sequence  
CC represents a HCV hydrophobic domain epitope which is used to illustrate  
CC the method of the invention  
XX  
SQ Sequence 18 AA;

Query Match 34.5%; Score 39; DB 3; Length 18;  
Best Local Similarity 61.5%; Pred. No. 58;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 SGVYPIDDDDYAS 16  
| | | | |  
Db 5 SNCYPVDPDYAS 17

Search completed: April 1, 2005, 21:47:11  
Job time : 177 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2005, 20:56:21 ; Search time 43 Seconds  
(without alignments)  
49.227 Million cell updates/sec

Title: US-10-776-989-9\_COPY\_20\_41  
Perfect score: 113  
Sequence: 1 EASGVYPIDDDDYASAGSGA 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 6283

Minimum DB seq length: 0  
Maximum DB seq length: 30

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : PIR 79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	33.6	8	A61328	trypsin (EC 3.4.21
2	32	28.3	14	PS0252	16K protein S404 -
3	31	27.4	20	S50022	trypsin-like prote
4	31	27.4	22	A34928	myosin heavy chain
5	30	26.5	19	B28457	proteoglycan II, b
6	30	26.5	29	A41683	hyaluronate recept
7	29	25.7	20	S48746	beta tubulin b-bet
8	29	25.7	25	S49025	ribosomal protein
9	28	24.8	13	A46463	preabsorbing antig
10	28	24.8	20	S50743	proteinase inhibit
11	28	24.8	24	161491	seed protein ws-9
12	28	24.8	27	PH1914	T-cell receptor be
13	27	23.9	11	S00616	paraspinal crystal
14	27	23.9	11	PT0301	Ig heavy chain CRD
15	27	23.9	21	PT0227	Ig heavy chain CDR
16	27	23.9	30	B60791	cytochrome P450c27
17	26	23.0	16	B42324	toxin II.6 - scorp
18	26	23.0	16	PH1475	T-cell receptor be
19	26	23.0	23	A32473	histidine-rich pro
20	26	23.0	26	S12146	hypothetical prote
21	25.5	22.6	25	A60067	neuromedin U-25 -
22	25	22.1	14	A39703	tubulin beta-3 cha
23	25	22.1	20	PH3460	alliin lyase (EC 4
24	25	22.1	23	PH1691	Ig heavy chain V r
25	25	22.1	25	S27229	prothymosin alpha
26	25	22.1	27	PL0151	nitrogen regulatio
27	25	22.1	27	S55234	flavodoxin isoform
28	25	22.1	28	PH0247	T-cell receptor Vb
29	24.5	21.7	27	S07646	hypothetical prote

30	24	21.2	6	2	JN0861	peptidyl-dipeptida
31	24	21.2	15	2	D54226	light-harvesting p
32	24	21.2	16	1	MTDFBS	melanotropin beta
33	24	21.2	17	2	S28839	RNA-binding protei
34	24	21.2	18	2	A39040	calsequestrin, car
35	24	21.2	18	4	I39461	anti-angiotensin,
36	24	21.2	20	2	A54042	mannose-6-phosphat
37	24	21.2	21	2	A27719	trypsin (EC 3.4.21
38	24	21.2	21	2	S58431	phosphatidylinosit
39	24	21.2	22	2	S13899	glycoprotein, 11.7
40	24	21.2	24	2	F45357	Kex2/subtilisin-li
41	24	21.2	24	2	B53524	ubiquinol-cytochro
42	24	21.2	26	2	A34280	phospholipase A2 (
43	24	21.2	26	2	PC1320	large granule l10
44	24	21.2	28	2	A56499	brevicidin-27 - Lact
45	24	21.2	30	2	C61125	glucagon-like pept
46	24	21.2	30	2	B61125	glucagon-like pept
47	24	21.2	30	2	B47607	immunogenic protei
48	23.5	20.8	20	2	A0198	N-acetylneuraminat
49	23.5	20.8	22	2	A53794	tyrocidine synthas
50	23	20.4	7	2	PT0623	T-cell receptor be
51	23	20.4	9	2	PH0918	T-cell receptor be
52	23	20.4	10	2	S13224	virG protein - Agr
53	23	20.4	10	2	G60787	sperm-activating p
54	23	20.4	10	2	E60787	sperm-activating p
55	23	20.4	12	2	S56122	type I DNA methylt
56	23	20.4	13	2	A26999	carboxylesterase (
57	23	20.4	15	2	PH1314	Ig heavy chain DJ
58	23	20.4	15	2	C44101	calmodulin, vasoac
59	23	20.4	15	2	B48047	phospholipase C-be
60	23	20.4	16	2	B36300	T-cell receptor de
61	23	20.4	16	2	PH1480	T-cell receptor be
62	23	20.4	18	2	PH0768	T-cell receptor be
63	23	20.4	20	2	C49753	hypothetical prote
64	23	20.4	21	2	G64564	hypothetical prote
65	23	20.4	21	2	H71858	hypothetical prote
66	23	20.4	22	2	A35418	brain natriuretic
67	23	20.4	22	2	S52357	hypothetical prote
68	23	20.4	23	2	A59048	convulsant peptide
69	23	20.4	23	2	A83698	hypothetical prote
70	23	20.4	24	2	S03455	T-cell receptor al
71	23	20.4	25	2	A48543	neuromedin U - chi
72	23	20.4	26	2	S78414	ribosomal protein
73	23	20.4	26	2	AG0756	hypothetical prote
74	23	20.4	27	2	PH1355	Ig heavy chain DJ
75	23	20.4	27	2	A61412	methane monooxygen
76	23	20.4	28	2	D47719	T-cell receptor al
77	23	20.4	28	2	D41912	T-cell receptor be
78	23	20.4	30	2	S29178	high-molecular-we
79	22.5	19.9	29	2	PQ0486	globulin 2a - taro
80	22	19.5	6	2	PT0593	T-cell receptor be
81	22	19.5	7	2	PT0611	T-cell receptor be
82	22	19.5	8	2	PT0595	T-cell receptor be
83	22	19.5	10	2	S70336	napin small chain
84	22	19.5	11	2	A58502	38K kidney stone p
85	22	19.5	11	2	PH0903	T-cell receptor be
86	22	19.5	11	2	S70338	napin small chain
87	22	19.5	12	2	PH1171	T-cell receptor al
88	22	19.5	13	2	PC4391	cysteine proteinas
89	22	19.5	15	2	S28515	T-cell receptor al
90	22	19.5	15	2	PS0452	32K protein 3306 -
91	22	19.5	15	2	PH0808	T-cell receptor al
92	22	19.5	16	2	A47393	neuropeptide calla
93	22	19.5	17	2	B49404	T-cell receptor be
94	22	19.5	17	2	S16929	flavodoxin A - Azo
95	22	19.5	18	2	B48839	T-cell receptor be
96	22	19.5	18	2	S28424	34K protein - rape
97	22	19.5	19	2	C40634	orf21 5' of eryK -
98	22	19.5	20	2	D34817	collagenolytic pro
99	22	19.5	20	2	F49048	T-cell receptor be
100	22	19.5	20	2	S03720	2S albumin small c



## ALIGNMENTS

```

RESULT 1
A61328
trypsin (EC 3.4.21.4) precursor - minke whale (fragment)
C:Species: Balaenoptera acutorostrata (minke whale, lesser rorqual)
C>Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 09-Jul-2004
C:Accession: A61328
R:Bricteux-Gregoire, S.; Schyns, R.; Florkin, M.; Emmens, M.; Welling, G.W.; Beintema, J.
Biochim. Biophys. Acta 386, 244-255, 1975
A:Title: N-terminal amino acid sequence of trypsinogen from the lesser rorqual, Balaenop
creas.
A:Reference number: A61328; MUID:75146765; PMID:1125273
A:Accession: A61328
A:Molecule type: protein
A:Residues: 1-8 <BRI>
A:Cross-references: UNIPROT:Q7M390
C:Keywords: hydrolase; protein digestion; serine proteinase; zymogen
F;1-8/Domain: activation peptide #status experimental <APT>

Query Match 33.6%; Score 38; DB 2; Length 8;
Best Local Similarity 85.7%; Pred. No. 2.8e+05;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 YPIDDDDD 13
DB 1 FPIDDDDD 7

RESULT 2
PS0252
16K protein 5404 - rice (strain Nihonbare) (fragment)
C:Species: Oryza sativa (rice)
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 05-Apr-1995
C:Accession: PS0252
R:Tsugita, A.
submitted to JIPID, April 1993
A:Reference number: PS0206
A:Accession: PS0252
A:Molecule type: protein
A:Residues: 1-14 <TSU>
A:Experimental source: strain Nihonbare
C:Comment: Molecular weight 16K, pI 4.9.

Query Match 28.3%; Score 32; DB 2; Length 14;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 6 VYPIDDDDY 14
DB 2 VLEADDDDY 10

RESULT 3
S50022
trypsin-like protein - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 07-May-1999
C:Accession: S50022
R:Koshikawa, N.; Yasumitsu, H.; Nagashima, Y.; Umeda, M.; Miyazaki, K.
Biochem. J. 303, 187-190, 1994
A:Title: Identification of one- and two-chain forms of trypsinogen 1 produced by a human
A:Reference number: S50020; MUID:95031957; PMID:7945238
A:Accession: S50022
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <KOS>
C:Superfamily: trypsin; trypsin homology

Query Match 27.4%; Score 31; DB 2; Length 20;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 8 PIDDDDD 13
DB 2 PFDDDD 7

```

## RESULT 4

```

A34928
myosin heavy chain, brain - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C>Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 13-Feb-1998
C:Accession: A34928
R:Murakami, N.; Healy-Louie, G.; Elzinga, M.
J. Biol. Chem. 265, 1041-1047, 1990
A:Title: Amino acid sequence around the serine phosphorylated by casein kinase II in brai
A:Reference number: A34928; MUID:90110084; PMID:2104826
A:Accession: A34928
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-22 <MUR>
C:Superfamily: myosin heavy chain; myosin motor domain homology

Query Match 27.4%; Score 31; DB 2; Length 22;
Best Local Similarity 66.7%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 10 DDDDYASAS 18
DB 5 DDDDESAS 13

```

## RESULT 5

```

B28457
proteoglycan II, bone - human (tentative sequence) (fragment)
C:Species: Homo sapiens (man)
C>Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 09-Jul-2004
C:Accession: B28457
R:Fisher, L.W.; Hawkins, G.R.; Tuross, N.; Termine, J.D.
J. Biol. Chem. 262, 9702-9708, 1987
A:Title: Purification and partial characterization of small proteoglycans I and II, bone
A:Reference number: A92656; MUID:87250639; PMID:3597437
A:Accession: B28457
A:Molecule type: protein
A:Residues: 1-19 <FIS>
A:Cross-references: UNIPROT:P07585

Query Match 26.5%; Score 30; DB 2; Length 19;
Best Local Similarity 46.2%; Pred. No. 3.4e+02;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 1 BEASGVYPIDDDDD 13
DB 1 DEAXGIAPEVPDD 13

```

## RESULT 6

```

A41683
hyaluronate receptor CD44 (version 2) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 30-Sep-1993
C:Accession: A41683
R:Shivelman, E.; Bishop, J.M.
Mol. Cell. Biol. 11, 5446-5453, 1991
A:Title: Expression of CD44 is repressed in neuroblastoma cells.
A:Reference number: A41683; MUID:92017823; PMID:1922057
A:Accession: A41683
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-29 <SHT>
A:Cross-references: GB:M69215
C:Keywords: cell adhesion

Query Match 26.5%; Score 30; DB 2; Length 29;

```

Best Local Similarity 60.0%; Pred. No. 5.4e+02; Mismatches 3; Indels 0; Gaps 0;  
Matches 6; Conservative 1;  
QY 11 DDDYASASGS 20  
Db 1 DDDVSSGSS 10  
RESULT 7  
S48746  
beta tubulin b-beta-2 - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 09-Jul-2004  
C:Accession: S48746  
R:Mary, J.; Redeker, V.; le Caer, J.P.; Prome, J.C.; Rossier, J.  
FEBS Lett. 353, 89-94, 1994  
A:Title: Class I and IVa beta-tubulin isoforms expressed in adult mouse brain are glutamylated  
A:Reference number: S48741; MUID:95010772; PMID:7326030  
A:Accession: S48746  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-20 <MAR>  
A:Cross-references: UNIPROT:Q9CRT0  
Query Match 25.7%; Score 29; DB 2; Length 20;  
Best Local Similarity 36.8%; Pred. No. 5.1e+02; Mismatches 6; Indels 2; Gaps 1;  
Matches 7; Conservative 4;  
QY 1 EASGVYPDDDDYASAG 19  
Db 2 EEEGEMV--EDDESEAAQ 18  
RESULT 8  
S49025  
ribosomal protein S12 [validated] - Haloarcula marismortui (fragment)  
C:Species: Haloarcula marismortui  
C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 12-Jun-2003  
C:Accession: S63966; S49025  
R:Engemann, S.; Noelle, R.; Herfurth, E.; Briesemeister, U.; Grelle, G.; Wittmann-Liebold, J.; Blochem, 234, 24-31, 1995  
A:Title: Cartography of ribosomal proteins of the 30S subunit from the halophilic Haloarcula marismortui  
A:Reference number: S63964; MUID:96096717; PMID:8529646  
A:Accession: S63966  
A:Molecule type: protein  
A:Residues: 1-25 <EN2>  
A:Cross-references: UNIPROT:Q7MS40  
C:Superfamily: ribosomal protein S12  
C:Keywords: protein biosynthesis; ribosome  
Query Match 25.7%; Score 29; DB 2; Length 25;  
Best Local Similarity 100.0%; Pred. No. 6.5e+02; Mismatches 0; Indels 0; Gaps 0;  
Matches 5; Conservative 0;  
QY 11 DDDYA 15  
Db 21 DDDYA 25  
RESULT 9  
A46463  
preabsorbing antigen (PA-Ag) - Streptococcus sp. (fragment)  
C:Species: Streptococcus sp.  
C>Date: 18-Jun-1993 #sequence\_revision 25-Apr-1997 #text\_change 20-Sep-1999  
C:Accession: A46463  
R:Yoshizawa, N.; Oshima, S.; Sagel, I.; Shimizu, J.; Treiser, G.  
J. Immunol. 148, 3110-3116, 1992  
A:Title: Role of a streptococcal antigen in the pathogenesis of acute poststreptococcal glomerulonephritis  
A:Reference number: A46463; MUID:92251171; PMID:1578137  
A:Accession: A46463  
A:Molecule type: protein  
A:Residues: 1-13 <YOS>  
A:Cross-references: PIDN:AAB22148.1; PID:G249161

A>Note: sequence extracted from NCBI backbone (NCBIP:100561)  
Query Match 24.8%; Score 28; DB 2; Length 13;  
Best Local Similarity 54.5%; Pred. No. 4.4e+02; Mismatches 1; Indels 0; Gaps 0;  
Matches 6; Conservative 1;  
QY 11 DDDYASASGS 21  
Db 2 DPEAASAFGDG 12  
RESULT 10  
S50743  
proteinase inhibitor (Bowman-Birk) C-II-related protein - potato (fragment)  
N:Alteinate names: 10K protein  
C:Species: Solanum tuberosum (potato)  
C>Date: 14-Jul-1995 #sequence\_revision 12-Apr-1996 #text\_change 09-Jul-2004  
C:Accession: S50743  
R:Mitsumori, C.; Yamagishi, K.; Fujino, K.; Kikuta, Y.  
Plant Mol. Biol. 26, 961-969, 1994  
A:Title: Detection of immunologically related Kunitz and Bowman-Birk proteinase inhibitors  
A:Reference number: S50741; MUID:95093035; PMID:8000008  
A:Accession: S50743  
A:Molecule type: protein  
A:Residues: 1-20 <MIT>  
A:Cross-references: UNIPROT:Q9S8K0  
Query Match 24.8%; Score 28; DB 2; Length 20;  
Best Local Similarity 45.5%; Pred. No. 7.1e+02; Mismatches 3; Indels 0; Gaps 0;  
Matches 5; Conservative 3;  
QY 8 PIDDDYASAS 18  
Db 7 PCDDEEATKS 17  
RESULT 11  
I61491  
seed protein ws-9 - winged bean (fragment)  
C:Species: Psophocarpus tetragonolobus (winged bean)  
C>Date: 07-Oct-1994 #sequence\_revision 07-Oct-1994 #text\_change 09-Jul-2004  
C:Accession: I61491  
R:Hirano, H.  
J. Protein Chem. 8, 115-130, 1989  
A:Title: Microsequence analysis of winged bean seed proteins electrophoretically separated from two-dimensional gel electrophoresis  
A:Reference number: A61491; MUID:89351606; PMID:2765119  
A:Accession: I61491  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-24 <HIR>  
A:Cross-references: UNIPROT:Q43708  
C:Superfamily: plant Kunitz-type proteinase inhibitor  
C:Keywords: seed  
Query Match 24.8%; Score 28; DB 2; Length 24;  
Best Local Similarity 50.0%; Pred. No. 8.7e+02; Mismatches 1; Indels 0; Gaps 0;  
Matches 5; Conservative 1;  
QY 11 DDDYASASGS 20  
Db 1 DDDLVDAGN 10  
RESULT 12  
PH1914  
T-cell receptor beta chain (clone A96) - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 10-Mar-1994 #sequence\_revision 28-Oct-1994 #text\_change 17-Mar-1999  
C:Accession: PH1914  
R:Sensai, M.; Salvi, S.; Castelli, C.; Maccalli, C.; Mazzocchi, A.; Mortarini, R.; Nicolini, J. Exp. Med. 178, 1231-1246, 1993  
A:Title: T cell receptor (TCR) structure of autologous melanoma-reactive cytotoxic T lymphocyte clone  
A:Reference number: PH1914; MUID:95093035; PMID:8000008

A;Reference number: PH1907; MUID:93389388; PMID:8376931

A;Accession: PH1914

A;Molecule type: mRNA

A;Residues: 1-27 <SEN>

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: T-cell receptor

Query Match 24.8%; Score 28; DB 2; Length 27;  
Best Local Similarity 62.5%; Pred. No. 9.9e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 14 YASASGSG 21

Db 5 YSAASGSG 12

#### RESULT 13

S00616

N;Alternate names: delta-endotoxin; parasporal crystal protein positive chain

C;Species: *Bacillus thuringiensis*

C;Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 09-Jul-2004

C;Accession: S00616

R;Chesukhina, G.G.; Kostina, L.I.; Zalunin, I.A.; Khodova, O.M.; Stepanov, V.M.

FEBS Lett. 232, 249-251, 1988

A;Title: *Bacillus thuringiensis* ssp. *galleriae* simultaneously produces two delta-endotox

A;Reference number: S00615

A;Accession: S00616

A;Molecule type: protein

A;Residues: 1-11 <CHE>

A;Cross-references: UNIPROT:Q7M154

C;Comment: This toxin is effective against the larvae of *Galleria mellonella* (greater wax

C;Superfamily: parasporal crystal protein

C;Keywords: delta-endotoxin

Query Match 23.9%; Score 27; DB 2; Length 11;  
Best Local Similarity 36.4%; Pred. No. 5.2e+02;  
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 7 YPIDDDDYASA 17

Db 1 YPLANNPYSGA 11

#### RESULT 14

PT0301

Ig heavy chain CDR3 region (clone 5-109) - human (fragment)

C;Species: *Homo sapiens* (man)

C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996

C;Accession: PT0301

R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.

J. Exp. Med. 173, 395-407, 1991

A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j

A;Reference number: PT0222; MUID:91108337; PMID:1899102

A;Accession: PT0301

A;Molecule type: DNA

A;Residues: 1-11 <YAM>

A;Experimental source: B lymphocyte

C;Keywords: heterotetramer; immunoglobulin

Query Match 23.9%; Score 27; DB 2; Length 11;  
Best Local Similarity 80.0%; Pred. No. 5.2e+02;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SGVYP 8

Db 4 SGYP 8

#### RESULT 15

PT0227

Ig heavy chain CDR3 region (clone 1-106) - human (fragment)

C;Species: *Homo sapiens* (man)

C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Aug-1996

C;Accession: PT0227

R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.

J. Exp. Med. 173, 395-407, 1991

A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j

A;Reference number: PT0222; MUID:91108337; PMID:1899102

A;Accession: PT0227

A;Molecule type: DNA

A;Residues: 1-21 <YAM>

A;Experimental source: B lymphocyte

C;Keywords: heterotetramer; immunoglobulin

Query Match 23.9%; Score 27; DB 2; Length 21;  
Best Local Similarity 62.5%; Pred. No. 1.1e+03;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 SGVYPIDD 11

Db 8 SGVYPLFD 15

#### RESULT 16

B60791

toxin II.6 - scorpion (*Centruroides limpidus*) (fragment)

C;Species: *Centruroides limpidus* limpidus

C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004

C;Accession: B60791

R;Alagon, A.C.; Guzman, H.S.; Martin, B.M.; Ramirez, A.N.; Carbone, E.; Possani, L.D.

Comp. Biochem. Physiol. B 89, 153-161, 1988

A;Title: Isolation and characterization of two toxins from the Mexican scorpion *Centruroides*

A;Reference number: A60791; MUID:88185033; PMID:2451580

A;Accession: B60791

A;Molecule type: protein

A;Residues: 1-30 <ALA>

A;Cross-references: UNIPROT:P59899

C;Comment: This peptide is toxic to mice.

C;Superfamily: scorpion neurotoxin

C;Keywords: neurotoxin

Query Match 23.9%; Score 27; DB 2; Length 30;  
Best Local Similarity 50.0%; Pred. No. 1.6e+03;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 7 YPIDDDDD 14

Db 17 YKLGNDY 24

#### RESULT 17

B42324

cytochrome P450c27/25 - rat (fragment)

C;Species: *Rattus norvegicus* (Norway rat)

C;Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 05-Mar-1999

C;Accession: B42324

R;Shayig, R.M.; Avadhani, N.G.

J. Biol. Chem. 267, 2421-2428, 1992

A;Title: Sequence complementarity between the 5'-terminal regions of mRNAs for rat mitoch

ap.

A;Reference number: A42324; MUID:92129322; PMID:1733943

A;Accession: B42324

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-16 <SHA>

A;Note: sequence extracted from NCBI backbone (NCBIN:78408, NCBI:P:88990)

C;Superfamily: human cytochrome P450 CYP11B1; cytochrome P450 homology

C;Keywords: heme; transmembrane protein

Query Match 23.0%; Score 26; DB 2; Length 16;  
Best Local Similarity 71.4%; Pred. No. 1.1e+03;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 GVPIDD 11

Db 11

Db 9 GKYPIRD 15

RESULT 18  
PH1475  
T-cell receptor beta chain (clone 223/5) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C>Date: 10-Mar-1994 #sequence\_revision 10-Mar-1994 #text\_change 15-Mar-2004  
C:Accession: PH1475  
R:Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.; K...  
J. Exp. Med. 177, 811-820, 1993  
A:Title: T cell receptor selection by and recognition of two class I major histocompatib...  
A:Reference number: PH1430; MUID:93171821; PMID:8436911  
A:Accession: PH1475  
A:Molecule type: mRNA  
A:Residues: 1-16 <CAS>  
A:Experimental source: cytolytic T-lymphocyte  
C:Keywords: receptor; T-cell

Query Match 23.0%; Score 26; DB 2; Length 16;  
Best Local Similarity 71.4%; Pred. No. 1.1e+03;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 15 ASASGSG 21  
||:|:|:  
Db 2 ASSSGTG 8

RESULT 19  
A32473  
histidine-rich protein C - liver fluke (fragment)  
C:Species: Fasciola hepatica (liver fluke)  
C>Date: 25-Sep-1989 #sequence\_revision 12-May-1994 #text\_change 09-Jul-2004  
C:Accession: A32473  
R:Waite, J.H.; Rice-Ficht, A.C.  
Biochemistry 28, 6104-6110, 1989  
A:Title: A histidine-rich protein from the vitellaria of the liver fluke Fasciola hepatic...  
A:Reference number: A32473; MUID:89375343; PMID:2775756  
A:Accession: A32473  
A:Molecule type: protein  
A:Residues: 1-23 <WAI>  
A:Cross-references: UNIPROT:Q7M3Q9  
A:Note: 22-Gly, 22-Ser, 23-Gly, and 23-Ser were also found  
C:Keywords: egg yolk  
F;6,8,12,16/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #status experimental

Query Match 23.0%; Score 26; DB 2; Length 23;  
Best Local Similarity 50.0%; Pred. No. 1.6e+03;  
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 10 DDDDYASASGSG 21  
| | | | |  
Db 4 DGYGSGYGS 15

RESULT 20  
S12146  
hypothetical protein E16 - phage D108 (fragment)  
C:Species: phage D108  
C>Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S12146  
R:Pato, M.L.; Banerjee, M.; Wagonner, B.T.  
Nucleic Acids Res. 18, 6458, 1990  
A:Title: Sequence of gene E15 of bacteriophage D108 and comparison with phage Mu.  
A:Reference number: S12145; MUID:91057162; PMID:2147059  
A:Accession: S12146  
A:Molecule type: DNA  
A:Residues: 1-26 <PAT>  
A:Cross-references: UNIPROT:P24796; EMBL:X54298; NID:g14794; PIDN:CAA38198.1; PID:g14796

Query Match 23.0%; Score 26; DB 2; Length 26;  
Best Local Similarity 66.7%; Pred. No. 1.9e+03;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 9 IDDDDY 14  
:|:|:|:  
Db 19 LDDDTY 24

RESULT 21  
A60067  
neuromedin U-25 - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C>Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 09-Jul-2004  
C:Accession: A60067  
R:Kage, R.; O'Harte, F.; Thim, L.; Conlon, J.M.  
Regul. Pept. 33, 191-198, 1991  
A:Title: Rabbit neuromedin U-25: lack of conservation of a posttranslational processing e...  
A:Reference number: A60067; MUID:91352345; PMID:1882085  
A:Accession: A60067  
A:Molecule type: protein  
A:Residues: 1-25 <KAG>  
A:Cross-references: UNIPROT:P34965  
C:Keywords: amidated carboxyl end; hormone; intestine  
F;25/Modified site: amidated carboxyl end (Asn) #status experimental

Query Match 22.6%; Score 25.5; DB 2; Length 25;  
Best Local Similarity 35.7%; Pred. No. 2.1e+03;  
Matches 5; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

QY 7 YPIDDDDYASASGS 20  
:|:|:|:|:  
Db 1 FPV-DEEFOSPFGS 13

RESULT 22  
A39703  
tubulin beta-3 chain - bovine (fragment)  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 08-Nov-1991 #sequence\_revision 08-Nov-1991 #text\_change 09-Jul-2004  
C:Accession: A39703  
R:Alexander, J.E.; Hunt, D.F.; Lee, M.K.; Shabanowitz, J.; Michel, H.; Berlin, S.C.; MacI...  
Proc. Natl. Acad. Sci. U.S.A. 88, 4685-4689, 1991  
A:Title: Characterization of posttranslational modifications in neuron-specific class III...  
A:Reference number: A39703; MUID:91271258; PMID:2052551  
A:Accession: A39703  
A:Molecule type: protein  
A:Status: preliminary  
A:Residues: 1-14 <ALE>  
A:Cross-references: UNIPROT:Q7M2L3

Query Match 22.1%; Score 25; DB 2; Length 14;  
Best Local Similarity 40.0%; Pred. No. 1.3e+03;  
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 10 DDDDYASASG 19  
| | | | |  
Db 3 DDEESEAOQ 12

RESULT 23  
S35460  
allin lyase (EC 4.4.1.4) - ramson (fragment)  
N:Alternate names: allinase  
C:Species: Allium ursinum (ramson)  
C>Date: 06-Feb-1995 #sequence\_revision 06-Feb-1995 #text\_change 09-Jul-2004  
C:Accession: S35460  
R:van Damme, E.J.M.; Smeets, K.; Torrekens, S.; van Leuven, F.; Peumans, W.J.  
Eur. J. Biochem. 209, 751-757, 1992  
A:Title: Isolation and characterization of allinase cDNA clones from garlic (Allium sativ...  
A:Reference number: S29300; MUID:93049322; PMID:1385120  
A:Accession: S35460  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-20 <DAM>

A;Cross-references: UNIPROT:Q7M1T5  
C;Keywords: carbon-sulfur lyase

Query Match 22.1%; Score 25; DB 2; Length 20;  
Best Local Similarity 60.0%; Pred. No. 2e+03;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 EASGVYPID 10  
||| |  
Db 10 EEAFAVAID 19

## RESULT 24

PH1691  
IG heavy chain V region (clone NP-7-4) - mouse (fragment)  
C;Species: Mus musculus (house mouse)  
C;Date: 24-Feb-1994 #sequence\_revision 24-Feb-1994 #text\_change 17-Mar-1999  
C;Accession: PH1691  
R;McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.  
J. Exp. Med. 178, 295-307, 1993  
A;Title: Antigen-driven B cell differentiation in vivo.  
A;Reference number: PH1675; MUID:93301607; PMID:8315385  
A;Accession: PH1691  
A;Molecule type: mRNA  
A;Residues: 1-23 <MCH>  
A;Experimental source: B cell  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin

Query Match 22.1%; Score 25; DB 2; Length 23;  
Best Local Similarity 71.4%; Pred. No. 2.3e+03;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 EASGVYP 8  
||| |  
Db 12 EARGNYP 18

## RESULT 25

S27229  
prothymosin alpha homolog - Escherichia coli (fragments)  
C;Species: Escherichia coli  
C;Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 10-Nov-1995  
C;Accession: S27229  
R;Vartapetian, A.; Chichkova, N.; Lyakhov, I.; Makarova, T.; Evstafieva, A.; Bogdanov, A.  
FEBS Lett. 313, 95-97, 1992  
A;Title: Segments of Escherichia coli genome similar to the exons of human prothymosin a  
A;Reference number: S27229; MUID:93050251; PMID:1426289  
A;Accession: S27229  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-6;7-25 <VAR>  
A;Note: the authors did not translate the codon for residues 17

Query Match 22.1%; Score 25; DB 2; Length 25;  
Best Local Similarity 41.7%; Pred. No. 2.5e+03;  
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 EASGVPIDDDD 13  
| | | |  
Db 8 ENGELEADNDD 19

Search completed: April 1, 2005, 21:51:03  
Job time : 46 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2005, 20:54:05 ; Search time 178 Seconds  
(without alignments)  
63.291 Million cell updates/sec

Title: US-10-776-989-9\_COPY\_20\_41  
Perfect score: 113  
Sequence: 1 BEASGVYPIDDDYASAGSGA 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 26236

Minimum DB seq length: 0  
Maximum DB seq length: 30

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : Uniprot\_03: \*  
1: uniprot\_sprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	38	33.6	8	2	Q7M390
2	38	33.6	16	1	TRYP_FELCA
3	37	32.7	18	2	Q9QW82
4	35	31.0	29	2	Q9PTV5
5	34	30.1	24	2	Q6KY7
6	33	29.2	29	2	Q6LD43
7	32	28.3	25	2	O11900
8	32	28.3	25	2	O11902
9	31	27.4	17	2	Q8I9U6
10	31	27.4	22	2	Q72445
11	31	27.4	23	2	Q702X4
12	31	27.4	30	2	Q711W7
13	30	26.5	17	2	Q6LDM6
14	30	26.5	25	2	Q9SKQ9
15	29	25.7	25	2	Q7M540
16	29	25.7	28	2	Q9QX35
17	28.5	25.2	21	2	Q9N0F8
18	28	24.8	10	2	P96423
19	28	24.8	13	2	Q9R5L4
20	28	24.8	15	2	Q40562
21	28	24.8	16	2	Q95Y34
22	28	24.8	19	1	AL22_HORSE
23	28	24.8	20	2	Q9S8K0
24	28	24.8	21	2	Q6S2K4
25	28	24.8	23	2	Q70ZU8
26	28	24.8	24	2	Q6S6S9
27	28	24.8	27	2	Q31210
28	28	24.8	29	1	AL21_HORSE
29	28	24.8	29	2	Q9UN87
30	28	24.8	29	2	Q711W3
31	28	24.8	29	2	Q711W5

32	28	24.8	30	2	Q711W9	Q711W9 bradyrhizob
33	27	23.9	11	2	Q7M154	Q7M154 bacillus th
34	27	23.9	19	2	Q70ZM0	Q70ZM0 bradyrhizob
35	27	23.9	20	1	GBB4_RAT	O35353 rattus norv
36	27	23.9	20	1	MEFP_MYTED	P83148 mytilus edu
37	27	23.9	20	2	Q711V5	Q711V5 rhizobium s
38	27	23.9	21	2	Q16218	Q16218 homo sapien
39	27	23.9	22	2	Q7RIQ0	Q7RIQ0 plasmodium
40	27	23.9	23	2	Q7S8C7	Q7S8C7 neurospora
41	27	23.9	24	2	Q8CL66	Q8CL66 yersinia pe
42	27	23.9	25	2	Q9BM25	Q9BM25 echinometra
43	27	23.9	25	2	Q9BM26	Q9BM26 echinometra
44	27	23.9	26	2	Q9NFH6	Q9NFH6 plasmodium
45	27	23.9	26	2	O56500	O56500 hepatitis c
46	27	23.9	26	2	O56516	O56516 hepatitis c
47	27	23.9	26	2	O56517	O56517 hepatitis c
48	27	23.9	26	2	O56518	O56518 hepatitis c
49	27	23.9	26	2	O56519	O56519 hepatitis c
50	27	23.9	27	2	Q16176	Q16176 drosophila
51	27	23.9	27	2	O56467	O56467 hepatitis c
52	27	23.9	27	2	O56492	O56492 hepatitis c
53	27	23.9	27	2	O56493	O56493 hepatitis c
54	27	23.9	27	2	O56513	O56513 hepatitis c
55	27	23.9	27	2	O56514	O56514 hepatitis c
56	27	23.9	27	2	O56529	O56529 hepatitis c
57	27	23.9	28	2	O95737	O95737 homo sapien
58	27	23.9	28	2	P70542	P70542 rattus norv
59	27	23.9	30	2	Q6WYF6	Q6WYF6 bacterioph
60	26.5	23.5	17	2	Q9UMN3	Q9UMN3 homo sapien
61	26.5	23.5	17	2	Q90737	Q90737 gallus gall
62	26.5	23.5	17	2	Q66269	Q66269 cucurbit mo
63	26	23.0	14	2	P96347	P96347 helicobacte
64	26	23.0	15	1	DIDH_PSESP	P80701 pseudomonas
65	26	23.0	15	2	O40563	O40563 nicotiana t
66	26	23.0	15	2	P83331	P83331 streptococ
67	26	23.0	18	2	Q9QUX0	Q9QUX0 mus sp. t
68	26	23.0	20	2	Q9URC7	Q9URC7 saccharomyc
69	26	23.0	21	2	Q6W8P8	Q6W8P8 homo sapien
70	26	23.0	22	2	Q9UCT4	Q9UCT4 homo sapien
71	26	23.0	23	2	Q7M3Q9	Q7M3Q9 fasciola he
72	26	23.0	26	1	E16_BPD10	P24796 bacterioph
73	26	23.0	27	2	O56515	O56515 hepatitis c
74	25.5	22.6	18	1	SODM_MYCHA	P80582 mycobacteri
75	25.5	22.6	25	1	NEOU_RABIT	P34965 mycobacteri
76	25	22.1	10	2	P96305	P96305 alteromonas
77	25	22.1	14	2	Q7M2L3	Q7M2L3 bos taurus
78	25	22.1	16	2	Q9R5S9	Q9R5S9 treponema d
79	25	22.1	17	2	Q9J131	Q9J131 glycine max
80	25	22.1	17	2	Q84450	Q84450 mus musculu
81	25	22.1	20	2	Q70Z46	Q70Z46 sus scrofa
82	25	22.1	20	2	Q7M1T5	Q7M1T5 allium ursi
83	25	22.1	22	2	Q9UWK5	Q9UWK5 methanobact
84	25	22.1	23	2	Q8C5M9	Q8C5M9 mus musculu
85	25	22.1	25	2	Q9SKR0	Q9SKR0 sus scrofa
86	25	22.1	25	2	Q940X4	Q940X4 glycine max
87	25	22.1	25	2	Q66RM9	Q66RM9 hevea bras
88	25	22.1	27	1	NTRC_VIBAL	P19905 vibrio algi
89	25	22.1	27	2	Q9S8L3	Q9S8L3 chlorella f
90	25	22.1	28	2	Q97144	Q97144 staphylococ
91	25	22.1	28	2	Q9R5C7	Q9R5C7 vibrio para
92	25	22.1	28	2	Q8CFB9	Q8CFB9 mus musculu
93	25	22.1	29	2	Q94JL6	Q94JL6 hevea bras
94	25	22.1	30	1	TL16_SPIOL	P81834 spinacia ol
95	25	22.1	30	2	Q9NFH4	Q9NFH4 plasmodium
96	25	22.1	30	2	Q8FV23	Q8FV23 brucella su
97	25	22.1	30	2	Q923S3	Q923S3 mus musculu
98	25	22.1	30	2	Q9QV43	Q9QV43 rattus sp.
99	25	22.1	30	2	Q991Q0	Q991Q0 human immun
100	24.5	21.7	21	1	FIBB_SYNCA	P14481 syncerus ca

ALIGNMENTS

```

Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 7 YPIDDDD 13
DB 1 FPIDDDD 7

RESULT 3
Q9QW82 PRELIMINARY; PRT; 18 AA.
AC Q9QW82;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE PHOSPHORIN-PEPTIDE fragment 12-45 (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN [1]_TaxID=10116;
RP SEQUENCE FROM N.A.
RX MEDLINE=91291127; PubMed=2064607;
RA Sabay B., Stetler-Stevenson W.G., Lechner J.H., Veis A.;
RT "Domain structure and sequence distribution in dentin phosphophoryn.";
RL Biochem. J. 276:699-707(1991).
FT NON_TER 1
FT NON_TER 18
SQ SEQUENCE 18 AA; 1993 MW; 8861C9E493B5A5BA CRC64;

Query Match 32.7%; Score 37; DB 2; Length 18;
Best Local Similarity 54.5%; Pred. No. 1.9e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 10 DDDYASASGS 20
DB 3 DDDYSDSDSS 13

RESULT 4
Q9PTV5 PRELIMINARY; PRT; 29 AA.
AC Q9PTV5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Low molecular mass polypeptide complex subunit 2 (Fragment).
GN Name=LMP2;
OS Oryzias latipes (Medaka fish) (Japanese ricefish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HNT;
RX MEDLINE=20082974; PubMed=10613850; DOI=10.1101/gr.9.12.1277;
RA Ohtsuka M., Makino S., Yoda K., Wada H., Naruse K., Mitani H.,
RA Shima A., Ozato K., Kimura M., Inoko H.;
RT "Construction of a linkage map of the medaka (Oryzias latipes) and
RT mapping of the Da mutant locus defective in dorsoventral patterning.";
RL Genome Res. 9:1277-1287(1999).
DR EMBL; AB030466; BAA85133.1; -
DR GO; GO:0005839; C:proteasome core complex (sensu Eukaryota); IEA.
DR GO; GO:0004175; F:endoropeptidase activity; IEA.
DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IEA.
DR InterPro; IPR001353; Proteasome_A_B.
FT NON_TER 1
FT NON_TER 29
SQ SEQUENCE 29 AA; 3061 MW; CA8D7FA59D645B95 CRC64;

Query Match 31.0%; Score 35; DB 2; Length 29;
Best Local Similarity 46.7%; Pred. No. 6.1e+02;
QY 7 YPIDDDD 13
DB 1 FPIDDDD 7

RESULT 1
Q7M390 PRELIMINARY; PRT; 8 AA.
AC Q7M390;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Trypsin (EC 3.4.21.4) precursor (Fragment).
OS Balaenoptera acutorostrata (Minke whale) (lesser rorqual).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Mysticeti;
OC Balaenopteridae; Balaenoptera.
NCBI_TaxID=9767;
RN [1]
RP SEQUENCE.
RA Briceux-Gregoire S., Schyns R., Florin M., Emmens M., Welling G.W.,
RA Beintema J.J.;
RT "N-terminal amino acid sequence of trypsinogen from the lesser
RT rorqual, Balaenoptera acutorostrata (Cetacea). Simultaneous isolation
RT of trypsinogen, chymotrypsinogen and ribonuclease from pancreas.";
RL Blochim. Biophys. Acta 386:244-255(1975).
DR PIR; A61328; A61328.
DR GO; GO:0004295; F:trypsin activity; IEA.
FT NON_TER 1
FT NON_TER 8
SQ SEQUENCE 8 AA; 964 MW; 8D3AAAAA04769 CRC64;

Query Match 33.6%; Score 38; DB 2; Length 8;
Best Local Similarity 85.7%; Pred. No. 1.6e+06;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 7 YPIDDDD 13
DB 1 FPIDDDD 7

RESULT 2
TRYP_FELCA STANDARD; PRT; 16 AA.
AC P81071;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Trypsin precursor (EC 3.4.21.4) (Fragment).
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
NCBI_TaxID=9685;
RN [1]
RP SEQUENCE.
RX TISSUE=Pancreas;
RX MEDLINE=97235546; PubMed=9080665; DOI=10.1016/S0305-0491(96)00214-3;
RA Steiner J.M., Medinger T.L., Williams D.A.;
RA "Purification and partial characterization of feline trypsin.";
RL Comp. Biochem. Physiol. 116B:87-93(1997).
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
CC -1- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: Belongs to the peptidase S1 family.
DR InterPro; IPR001254; Peptidase_S1.
DR PROSITE; PS00240; TRYPSIN_DOM; PARTIAL.
DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW Calcium-binding; Digestion; Direct protein sequencing; Hydrolase;
KW Pancreas; Serine protease; Zymogen.
FT PROPEP 1
FT CHAIN 9 >16 Activation peptide.
FT NON_TER 16
FT NON_TER 16
SQ SEQUENCE 16 AA; 1825 MW; A6D751B58760A86 CRC64;

Query Match 33.6%; Score 38; DB 1; Length 16;
Best Local Similarity 85.7%; Pred. No. 1.2e+02;

```





```

RESULT 9
ID Q819U6 PRELIMINARY; PRT; 17 AA.
AC Q819U6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Large conductance calcium activated potassium channel pSlo spliceform
DE 4C (Fragment).
DE Name=pslo;
GN Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattodea;
OC Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22557503; PubMed=12670308;
RA Derst C., Messutat S., Walther C., Eckert M., Heinemann S.H.,
RA Wicher D.;
RT "The large conductance Ca2+-activated potassium channel (pSlo) of the
RT cockroach Periplaneta americana: structure, localization in neurons
RT and electrophysiology."
RL Eur. J. Neurosci. 17:1197-1212(2003).
DR EMBL; AF452169; AAN76924.1; -.
DR GO; GO:0005216; F:ion channel activity; IEA.
KW Ionic channel.
KW NON_TER 1
FT NON_TER 17
SQ SEQUENCE 17 AA; 1975 MW; 3ED226217C9E173E CRC64;

Query Match 27.4%; Score 31; DB 2; Length 17;
Best Local Similarity 54.5%; Pred. No. 1.4e+03;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 11 DDYASASGSG 21
Db 4 DEDLYASEAG 14

RESULT 10
ID Q72445 PRELIMINARY; PRT; 22 AA.
AC Q72445;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Nonfunctional trypsin 1 (EC 3.4.21.4) (Fragment).
GN Names=PRSS1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22651503; PubMed=12765848; DOI=10.1016/S1096-7192(03)00050-7;
RA Chen J.-M., Le Marechal C., Lucas D., Raguenes O., Ferec C.;
RT "Loss of function" mutations in the cationic trypsinogen gene (PRSS1)
RT may act as a protective factor against pancreatitis."
RL Mol. Genet. Metab. 79:67-70(2003).
DR EMBL; AY254094; AAP42827.1; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004295; F:trypsin activity; IEA.
KW Hydrolase.
KW NON_TER 1
FT NON_TER 1
SQ SEQUENCE 22 AA; 2355 MW; 0F581AC1A31E9AC4 CRC64;

Query Match 27.4%; Score 31; DB 2; Length 22;
Best Local Similarity 83.3%; Pred. No. 1.8e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 8 PIDDDD 13
Db 3 PFDDDD 8

RESULT 11
ID Q70ZX4 PRELIMINARY; PRT; 23 AA.
AC Q70ZX4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Chitoooligosaccharide deacetylase (Fragment).
GN Name=nodB;
OS Bradyrhizobium sp. USDA 3152A.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=187031;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=USDA 3152A;
RA Moulin L., Bena G., Boivin-Masson C., Stepkowski T.;
RT "Phyl ogetic analyses of symbiotic nodulation genes support vertical
RT and lateral gen e co-transfer within the Bradyrhizobium genus."
RL Mol. Phylogenet. Evol. 0:0-0(2003).
DR EMBL; AJ430711; CAD23461.1; -.
DR NON_TER 23
SQ SEQUENCE 23 AA; 2527 MW; 620D19D1278DD828 CRC64;

Query Match 27.4%; Score 31; DB 2; Length 23;
Best Local Similarity 75.0%; Pred. No. 1.9e+03;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 13 DYASASGS 20
Db 13 DYADVSGS 20

RESULT 12
ID Q711W7 PRELIMINARY; PRT; 30 AA.
AC Q711W7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE NodB protein (Fragment).
GN Name=nodB;
OS Bradyrhizobium sp. ORS130.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=142633;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ORS130;
RA Moulin L.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AJ300254; CAC82882.1; -.
FT NON_TER 30
SQ SEQUENCE 30 AA; 3223 MW; C3367A880F1580C2 CRC64;

Query Match 27.4%; Score 31; DB 2; Length 30;
Best Local Similarity 75.0%; Pred. No. 2.5e+03;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 13 DYASASGS 20
Db 13 DYADVSGS 20

RESULT 13
ID Q6LDM6 PRELIMINARY; PRT; 17 AA.
AC Q6LDM6;

```

```
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Paromyosin (Fragment).
OS Onchocerca volvulus.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC Onchocercidae; Onchocerca.
OX NCBI_TaxID=6282;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90220759; PubMed=2325708; DOI=10.1016/0166-6851(90)90030-P;
RA Limberger R.J., McReynolds L.A.;
RT "Filarial paromyosin: cDNA sequences from Dirofilaria immitis and
RT Onchocerca volvulus."
RL Mol. Biochem. Parasitol. 38:271-280 (1990).
DR EMBL; M29740; AAA29429.1; -.
FT NON_TER 1
FT SEQUENCE 17 AA; 1909 MW; B86C55E8986B75A0 CRC64;
SQ SEQUENCE 17 AA; 1909 MW; B86C55E8986B75A0 CRC64;

Query Match 26.5%; Score 30; DB 2; Length 17;
Best Local Similarity 50.0%; Pred. No. 2e+03;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 ASGVPIIDD 12
DB |||:|:|
7 ASKIYVLEDD 16

RESULT 14
Q95KQ9 PRELIMINARY; PRT; 25 AA.
ID Q95KQ9
AC Q95KQ9
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cystatin B (Fragment).
GN Name=csb;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX Russo V., Fontanesi L., Davoli R., Nanni costa L., Cagnazzo M.,
RA Butazzoni L., Virgili R., Yerle M.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ315563; CAC85532.1; -.
DR HSP; P04080; I STF.
FT NON_TER 1
FT NON_TER 25
FT SEQUENCE 25 AA; 2950 MW; 87FBDDC4A882CBCA CRC64;
SQ SEQUENCE 25 AA; 2950 MW; 87FBDDC4A882CBCA CRC64;

Query Match 26.5%; Score 30; DB 2; Length 25;
Best Local Similarity 66.7%; Pred. No. 2.9e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 9 IDDDY 14
DB :|:|:
3 VDDDDP 8

RESULT 15
Q7M540 PRELIMINARY; PRT; 25 AA.
ID Q7M540
AC Q7M540
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ribosomal protein S12 (Fragment).
OS Halocaula marismortui (Halobacterium marismortui).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halocaula.
OX NCBI_TaxID=2238;
```

```
RN [1]
RP SEQUENCE.
RX MEDLINE=96096717; PubMed=8529646;
RA Engemann S., Noelle R., Herfurth E., Briesemeister U., Grelle G.,
RA Wittmann-Liebold B.;
RT "Cartography of ribosomal proteins of the 30S subunit from the
RT halophilic Halocaula marismortui and complete sequence analysis of
RT protein HS26."
RL Eur. J. Biochem. 234:24-31 (1995).
DR PIR; S63966; S63966.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005840; C:ribosome; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR006032; Ribosomal_S12_23.
DR ProDom; PD000576; Ribosomal_S12_23; 1.
FT NON_TER 1
FT NON_TER 25
FT SEQUENCE 25 AA; 3021 MW; CA03522877D8B52D CRC64;
SQ SEQUENCE 25 AA; 3021 MW; CA03522877D8B52D CRC64;

Query Match 25.7%; Score 29; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 4.2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 DDYA 15
DB |||:|
21 DDYA 25

RESULT 16
Q9QX35 PRELIMINARY; PRT; 28 AA.
ID Q9QX35
AC Q9QX35
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 20, Last annotation update)
DE Guanine nucleotide binding protein beta 2 subunit (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=98283588; PubMed=9622245; DOI=10.1016/S0306-4522(97)00623-4;
RA Betty M., Harnish S.W., Rhodes K.J., Cockett M.I.;
RT "Distribution of heterotrimeric G-protein beta and gamma subunits in
RT the rat brain."
RL Neurosci. 85:475-486 (1998).
DR EMBL; AF022084; AAB82551.1; -.
DR InterPro; IPR011044; Amine_DH_B_like.
DR InterPro; IPR001680; WD40.
DR ProDom; PD000018; WD40; 1.
DR PROSITE; PS50082; WD_REPEATS_2; 1.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW WD repeat.
FT NON_TER 1
FT SEQUENCE 28 AA; 3043 MW; 0F5639CC5CC388AE CRC64;
SQ SEQUENCE 28 AA; 3043 MW; 0F5639CC5CC388AE CRC64;

Query Match 25.7%; Score 29; DB 2; Length 28;
Best Local Similarity 50.0%; Pred. No. 4.7e+03;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 IDDDYASAGS 20
DB :|:|:|:|
8 VTDDGMVATGS 19

RESULT 17
Q9N0F8 PRELIMINARY; PRT; 21 AA.
ID Q9N0F8
AC Q9N0F8
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
```

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Class III beta-tubulin (Fragment).  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Arai K.;  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB022056; BAA96411.2; -;  
 FT NON\_TER 1  
 FT NON\_TER 21  
 SQ SEQUENCE 21 AA; 2401 MW; 1E7F67C713357D7 CRC64;  
  
 Query Match 25.28; Score 28.5; DB 2; Length 21;  
 Best Local Similarity 36.8%; Pred. No. 4.1e+03;  
 Matches 7; Conservative 4; Mismatches 7; Indels 1; Gaps 1;  
  
 QY 1 EEASGVVPIDDDVYASAG 19  
 || : | | | : | | |  
 Db 2 EEEGMYE-DDESEAAQ 19  
  
 RESULT 18  
 P96423 PRELIMINARY; PRT; 10 AA.  
 ID P96423; Q7DCH6;  
 AC P96423; Q7DCH6;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Coded portion of proteolysis tag (translated portion of tmRNA gene  
 DE ssa) (Fragment).  
 GN OrderedLocusNames=PA0826.1;  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Williams K.P., Bartel D.P.;  
 RL "Phylogenetic analysis of tmRNA secondary structure."; RNA 2:1306-1310(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;  
 RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 RT opportunistic pathogen."; Nature 406:959-964(2000).  
 RL EMBL; U68078; AAB48029.1; -;  
 DR EMBL; AB004517; AAC43616.1; -;  
 KW Complete proteome.  
 FT NON\_TER 1  
 FT NON\_TER 10  
 SQ SEQUENCE 10 AA; 1037 MW; 857BD22DCB544AAA CRC64;  
  
 Query Match 24.8%; Score 28; DB 2; Length 10;  
 Best Local Similarity 55.6%; Pred. No. 2.3e+03;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 10 DDDVYASAS 18  
 : | | | | |  
 Db 2 NDDNYALAA, 10

RESULT 19  
 Q9R5L4 PRELIMINARY; PRT; 13 AA.  
 ID Q9R5L4;  
 AC Q9R5L4;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE PREABSORBING antigen, PA-AG-STREPTOCOCCAL antigen (Fragment).  
 OS Streptococcus.  
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae.  
 OX NCBI\_TaxID=1301;  
 RN [1]  
 RP SEQUENCE.  
 RA YOSHIZAWA N., Oshima S., Sagel I., Shimizu J., Treser G.;  
 RL "Role of a streptococcal antigen in the pathogenesis of acute  
 RL poststreptococcal glomerulonephritis. Characterization of the antigen  
 RL and a proposed mechanism for the disease."; J. Immunol. 148:3110-3116(1992).  
 FT NON\_TER 1  
 FT NON\_TER 13  
 SQ SEQUENCE 13 AA; 1264 MW; CFF82EEC7D10CSBD CRC64;  
  
 Query Match 24.8%; Score 28; DB 2; Length 13;  
 Best Local Similarity 54.5%; Pred. No. 3e+03;  
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
  
 QY 11 DDDYASASGS 21  
 | : | | | |  
 Db 2 DPEAASAPGDG 12  
  
 RESULT 20  
 Q40562 PRELIMINARY; PRT; 15 AA.  
 ID Q40562;  
 AC Q40562;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Small ras-related protein (fragment).  
 GN Name=Ran-A2;  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Lamiales; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SR1;  
 RX MEDLINE=95078947; PubMed=7987414;  
 RA Merkle T., Haizel T., Matsumoto T., Harter K., Dallmann G., Nagy F.;  
 RT "Phenotype of the fission yeast cell cycle regulatory mutant pimi-46  
 RT is suppressed by a tobacco cDNA encoding a small, Ran-like GTP-binding  
 RT protein."; Plant J. 6:555-565(1994).  
 RL Plant J. 6:555-565(1994).  
 DR EMBL; L16786; AAA73564.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 15  
 SQ SEQUENCE 15 AA; 1707 MW; 8AB89F924EA30CAF CRC64;  
  
 Query Match 24.8%; Score 28; DB 2; Length 15;  
 Best Local Similarity 83.3%; Pred. No. 3.5e+03;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 8 PIDDDD 13  
 | | | | |  
 Db 6 PDDDDD 11  
  
 RESULT 21  
 Q95Y34 PRELIMINARY; PRT; 16 AA.  
 ID Q95Y34  
 AC Q95Y34;

DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Hypothetical protein Y108G3AL.6.  
GN Name=Y108G3AL.6; ORFNames=Y108G3AL.6;  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
OC Rhabditidae; Poloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RT "Genome sequence of the nematode C. elegans: a platform for  
RT investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Courtney L.;  
RT "The sequence of C. elegans cosmid Y108G3AL.";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.H.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
RN [8]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
RN [9]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
RN [10]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Wilson R.;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
RN [11]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Wilson R.;  
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
RN [12]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RG WormBase Consortium;  
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC024744; AAL13326.1; --  
DR WormBase; WBGene0002437; Y108G3AL.6.  
KW WormPep; Y108G3AL.6; CE29605.  
KW Hypothetical protein.  
SQ SEQUENCE 16 AA; 1904 MW; 4932099B5AA361E5 CRC64;  
  
Query Match 24.8%; Score 28; DB 2; Length 16;  
Best Local Similarity 66.7%; Pred. No. 3.7e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 9 IDDDY 14  
Db 3 VDDSDY 8  
  
RESULT 22  
AL22\_HORSE  
ID AL22\_HORSE STANDARD; PRT; 19 AA.  
AC P81217;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Dander allergen Equ c 2.0102 (Fragment).  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Skin;  
RX MEDLINE=98237590; PubMed=9578478;  
RA Bulone V., Krogstad-Johnsen T., Smetad-Paulsen B.;  
RT "Separation of horse dander allergen proteins by two-dimensional  
RT electrophoresis -- molecular characterization and identification of  
RT Equ c 2.0101 and Equ c 2.0102 as lipocalin proteins.";  
RL Eur. J. Biochem. 253:202-211(1998).  
RN [2]  
RP MASS SPECTROMETRY.  
RC TISSUE=Hair dandruff;  
RX MEDLINE=21257838; PubMed=11358533;  
RA Goubran Botros H., Poncet P., Rabillon J., Fontaine T., Laval J.-M.,  
RA David B.;  
RT "Biochemical characterization and surfactant properties of horse  
RT allergens.";  
RL Eur. J. Biochem. 268:3126-3136(2001).  
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).  
CC -!- MASS SPECTROMETRY: MW=16000; METHOD=Electrospray; RANGE=1-7;  
CC NOTE=Ref.2.  
CC -!- ALLERGEN: Causes an allergic reaction in human. Potent allergen of  
CC horse dander.  
CC -!- SIMILARITY: Belongs to the lipocalin family.  
CC -!- CAUTION: Equ c 2.0101 and c 2.0102 might be two variants of the  
CC same protein.  
DR InterPro; IPR000566; Lipocalin\_cytFABP.  
DR PROSITE; PS00213; LIPOCALIN; PARTIAL.  
KW Allergen; Direct protein sequencing; Lipocalin; Transport.  
FT NON\_TER 19  
SQ SEQUENCE 19 AA; 2195 MW; A8EE66FAFC9322C36 CRC64;  
  
Query Match 24.8%; Score 28; DB 1; Length 19;  
Best Local Similarity 41.7%; Pred. No. 4.4e+03;  
Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
  
QY 8 RIDDDDYASAG 19  
Db 4 POSDTYQLSG 15  
  
RESULT 23  
Q9S8K0  
ID Q9S8K0 PRELIMINARY; PRT; 20 AA.  
AC Q9S8K0;  
DT 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE 10 kDa C-II-like Bowman-Birk proteinase inhibitor (Fragment).  
OS Solanum tuberosum (potato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC lamiales; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4113;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=95093035; PubMed=8000008;  
RA Mitsumori C., Yamagishi K., Fujino K., Kikuta Y.;  
RT "Detection of immunologically related Kunitz and Bowman-Birk  
proteinase inhibitors expressed during potato tuber development.";  
RL Plant Mol. Biol. 26:961-969(1994).  
DR PIR; S50743; S50743.  
DR HSP; P01055; IHHI.  
SQ SEQUENCE 20 AA; 2135 MW; 18D5AA25C42BE53C CRC64;  
  
Query Match 24.8%; Score 28; DB 2; Length 20;  
Best Local Similarity 45.5%; Pred. No. 4.7e+03;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
  
QY 8 PIDDDYASAS 18  
| | | | | | | | | |  
Db 7 PCDDEEAATKS 17  
  
RESULT 24  
Q6SZK4 PRELIMINARY; PRT; 21 AA.  
AC Q6SZK4;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Hypothetical protein (fragment).  
OS Rhodococcus fascians.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Nocardiaceae; Rhodococcus.  
OX NCBI\_TaxID=1828;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NRRL-B-15096;  
RA Dorado G., Roldan J.M., Leshner J.;  
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY444551; AAR1271.1;  
KW Hypothetical protein.  
FT NON\_TER 1  
FT NON\_TER 21 21  
SQ SEQUENCE 21 AA; 2161 MW; B3032C3ED62A6108 CRC64;  
  
Query Match 24.8%; Score 28; DB 2; Length 21;  
Best Local Similarity 50.0%; Pred. No. 4.9e+03;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
  
QY 12 DDYASASGS 21  
| | | | | | | | | |  
Db 3 DDYGDPAFG 12  
  
RESULT 25  
Q70ZU8 PRELIMINARY; PRT; 23 AA.  
AC Q70ZU8;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Chitooligosaccharide deacetylase (fragment).  
GN Name=nodB;  
OS Bradyrhizobium sp. ORS1810.  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Bradyrhizobiaceae; Bradyrhizobium.  
OX NCBI\_TaxID=186897;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ORS1810;  
RA Moulin L., Bena G., Boivin-Masson C., Stepkowski T.;  
RT "Phylogenetic analyses of symbiotic nodulation genes support vertical  
and lateral gene co-transfer within the Bradyrhizobium genus.";  
RL Mol. Phylogenet. Evol. 0:0-0(2003).  
DR EMBL; AJ430724; CAD23487.1;  
FT NON\_TER 23 23  
SQ SEQUENCE 23 AA; 2584 MW; 72D7647CFD49C828 CRC64;  
  
Query Match 24.8%; Score 28; DB 2; Length 23;  
Best Local Similarity 62.5%; Pred. No. 5.4e+03;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 13 DYASASGS 20  
| | | | | | | | | |  
Db 13 DYADVTS 20

Search completed: April 1, 2005, 21:50:16  
Job time : 186 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2005, 21:35:41 ; Search time 42 Seconds  
(without alignments)  
39.102 Million cell updates/sec

Title: US-10-776-989-9\_COPY\_20\_41  
Perfect score: 113  
Sequence: 1 BRASGVPIDDDDYASAGSGA 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 233678

Minimum DB seq length: 0  
Maximum DB seq length: 30

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : Issued Patents.AA.\*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	44.2	19	3	US-09-169-015-52
2	50	44.2	19	4	US-09-626-581D-57
3	50	44.2	19	4	US-09-415-765B-57
4	50	44.2	19	4	US-09-626-580C-57
5	50	44.2	19	4	US-09-749-959-49
6	44	38.9	18	4	US-09-285-912A-136
7	41	36.3	14	1	US-09-740-876-1
8	41	36.3	14	2	US-08-723-726-1
9	41	36.3	16	3	US-08-390-874C-13
10	41	36.3	16	3	US-09-265-772-13
11	41	36.3	16	4	US-09-472-112-4
12	41	36.3	25	4	US-09-500-700-28
13	41	36.3	26	4	US-09-481-620A-34
14	41	36.3	27	3	US-08-672-213-39
15	41	36.3	27	4	US-09-781-804-19
16	39	34.5	14	3	US-08-632-514C-29
17	39	34.5	14	3	US-09-188-177-29
18	39	34.5	18	4	US-09-856-920-1
19	38	33.6	16	3	US-09-208-966-15
20	38	33.6	16	4	US-09-775-052A-15
21	37	32.7	12	3	US-08-194-560-3
22	37	32.7	13	3	US-08-895-495-40
23	37	32.7	18	3	US-09-467-997-9
24	37	32.7	23	1	US-08-507-124-18
25	37	32.7	23	1	US-08-507-124-33
26	37	32.7	25	1	US-08-634-060-27
27	37	32.7	25	2	US-08-700-846-10
28	36	31.9	10	1	US-08-219-242A-3
29	36	31.9	10	1	US-08-000-931-4
30	36	31.9	10	1	US-08-476-685D-3
31	36	31.9	10	1	US-08-276-852-5
32	36	31.9	10	1	US-08-285-935-13
33	36	31.9	10	1	US-08-472-877B-3
34	36	31.9	10	1	US-08-472-877B-10
35	36	31.9	10	1	US-08-899-575-5
36	36	31.9	10	1	US-08-487-860-13
37	36	31.9	10	1	US-08-899-575-5
38	36	31.9	10	2	US-08-357-642A-7
39	36	31.9	10	2	US-08-460-626-7
40	36	31.9	10	2	US-08-460-626-19
41	36	31.9	10	2	US-08-460-626-22
42	36	31.9	10	2	US-08-755-728-21
43	36	31.9	10	2	US-08-832-935-5
44	36	31.9	10	2	US-08-974-655-21
45	36	31.9	10	3	US-08-832-985-1
46	36	31.9	10	3	US-08-938-291A-10
47	36	31.9	10	3	US-09-072-958-3
48	36	31.9	10	3	US-09-283-011-21
49	36	31.9	10	3	US-09-081-345-13
50	36	31.9	10	3	US-08-861-105-12
51	36	31.9	10	3	US-09-054-918A-5
52	36	31.9	10	4	US-09-410-903-82
53	36	31.9	10	4	US-09-233-857-14
54	36	31.9	10	4	US-09-363-960-5
55	36	31.9	10	4	US-09-339-922A-28
56	36	31.9	10	4	US-08-835-159-1
57	36	31.9	10	4	US-09-589-619-10
58	36	31.9	10	4	US-08-791-391A-28
59	36	31.9	10	4	US-09-016-061-28
60	36	31.9	10	4	US-08-641-294-11
61	36	31.9	10	4	US-09-456-090A-30
62	36	31.9	10	4	US-09-012-135A-21
63	36	31.9	10	4	US-09-453-234-30
64	36	31.9	10	4	US-09-069-228-13
65	36	31.9	10	5	PCT-US95-08743-5
66	36	31.9	11	3	US-09-169-015-2
67	36	31.9	11	3	US-08-628-829-17
68	36	31.9	11	4	US-09-626-581D-2
69	36	31.9	11	4	US-09-415-765B-2
70	36	31.9	11	4	US-09-626-580C-2
71	36	31.9	11	4	US-09-285-912A-69
72	36	31.9	11	4	US-09-555-166-44
73	36	31.9	12	3	US-09-449-437A-11
74	36	31.9	12	3	US-09-058-483-5
75	36	31.9	13	3	US-08-157-562-2
76	36	31.9	13	5	PCT-US96-12374-4
77	36	31.9	15	1	US-08-321-071A-5
78	36	31.9	17	3	US-08-789-333F-51
79	36	31.9	17	4	US-08-787-738B-51
80	36	31.9	17	4	US-09-916-940-51
81	36	31.9	17	4	US-10-096-550-51
82	36	31.9	22	4	US-09-774-282-6
83	36	31.9	22	1	US-08-507-124-14
84	36	31.9	23	1	US-08-507-124-31
85	36	31.9	24	1	US-08-507-124-23
86	36	31.9	24	1	US-08-507-124-29
87	36	31.9	28	1	US-08-078-683A-15
88	36	31.9	28	3	US-08-861-105-6
89	36	31.9	28	4	US-08-471-970A-15
90	36	31.9	28	4	US-09-723-677B-15
91	35	31.0	24	1	US-08-507-124-17
92	35	31.0	24	1	US-08-507-124-21
93	35	31.0	24	1	US-08-507-124-24
94	35	31.0	24	1	US-08-507-124-26
95	35	31.0	24	1	US-08-507-124-27
96	35	31.0	24	1	US-08-507-124-30
97	35	31.0	24	1	US-09-235-230-12
98	34	30.1	11	4	US-09-374-964-1
99	34	30.1	21	1	US-09-297-269-21
100	34	30.1	21	4	US-09-639-207-14



ALIGNMENTS

RESULT 1  
US-09-169-015-52  
; Sequence 52, Application US/09169015  
; Patent No. 6180343  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David  
; TITLE OF INVENTION: Green Fluorescent Protein Fusions with Random Peptides  
; FILE REFERENCE: A66900/DJB/RMS/SJR  
; CURRENT APPLICATION NUMBER: US/09/169,015  
; CURRENT FILING DATE: 1998-10-08  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 52  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-169-015-52

Query Match 44.2%; Score 50; DB 3; Length 19;  
Best Local Similarity 58.8%; Pred. No. 0.5;  
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 GVPYDVPDYASLGSGG 21  
| | | | | | | | | |  
Db 3 GGYPTDVPDYASLGSGG 19

RESULT 2  
US-09-626-581D-57  
; Sequence 57, Application US/09626581D  
; Patent No. 6548249  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David  
; TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide  
; FILE REFERENCE: A-66900-3/RMS  
; CURRENT APPLICATION NUMBER: US/09/626,581D  
; CURRENT FILING DATE: 2000-07-27  
; PRIOR APPLICATION NUMBER: 09/169,015  
; PRIOR FILING DATE: 1998-10-08  
; PRIOR FILING DATE: 1999-10-08  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 57  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-626-581D-57

Query Match 44.2%; Score 50; DB 4; Length 19;  
Best Local Similarity 58.8%; Pred. No. 0.5;  
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 GVPYDVPDYASLGSGG 21  
| | | | | | | | | |  
Db 3 GGYPTDVPDYASLGSGG 19

RESULT 3  
US-09-415-765B-57  
; Sequence 57, Application US/09415765B  
; Patent No. 6548632

; GENERAL INFORMATION:  
; APPLICANT: Anderson, David  
; TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide  
; FILE REFERENCE: A66900-1/RMS/AMS  
; CURRENT APPLICATION NUMBER: US/09/415,765B  
; CURRENT FILING DATE: 1999-10-08  
; PRIOR APPLICATION NUMBER: 09/169,015  
; PRIOR FILING DATE: 1998-10-08  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 57  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
US-09-415-765B-57

Query Match 44.2%; Score 50; DB 4; Length 19;  
Best Local Similarity 58.8%; Pred. No. 0.5;  
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 GVPYDVPDYASLGSGG 21  
| | | | | | | | | |  
Db 3 GGYPTDVPDYASLGSGG 19

RESULT 4  
US-09-626-580C-57  
; Sequence 57, Application US/09626580C  
; Patent No. 6562617  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David  
; TITLE OF INVENTION: FUSIONS OF SCAFFOLD PROTEINS WITH RANDOM PEPTIDE LIBRARIES  
; FILE REFERENCE: A-66900-2/RMS/AMS  
; CURRENT APPLICATION NUMBER: US/09/626,580C  
; CURRENT FILING DATE: 2000-07-27  
; PRIOR APPLICATION NUMBER: US 09/415,765  
; PRIOR FILING DATE: 1999-10-08  
; PRIOR APPLICATION NUMBER: US 09/169,015  
; PRIOR FILING DATE: 1998-10-08  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 57  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-09-626-580C-57

Query Match 44.2%; Score 50; DB 4; Length 19;  
Best Local Similarity 58.8%; Pred. No. 0.5;  
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 GVPYDVPDYASLGSGG 21  
| | | | | | | | | |  
Db 3 GGYPTDVPDYASLGSGG 19

RESULT 5  
US-09-749-959-49  
; Sequence 49, Application US/09749959  
; Patent No. 6596485  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David  
; TITLE OF INVENTION: GREEN FLUORESCENT PROTEIN FUSIONS WITH RANDOM PEPTIDES  
; FILE REFERENCE: A-66900-5/RMS/AMS  
; CURRENT APPLICATION NUMBER: US/09/749,959

;/ CURRENT FILING DATE: 2000-12-27  
;/ PRIOR APPLICATION NUMBER: US 09/169,015  
;/ PRIOR FILING DATE: 1998-10-08  
;/ NUMBER OF SEQ ID NOS: 50  
;/ SOFTWARE: PatentIn version 3.1  
;/ SEQ ID NO 49  
;/ LENGTH: 19  
;/ TYPE: PRT  
;/ ORGANISM: Artificial sequence  
;/ FEATURE:  
;/ OTHER INFORMATION: synthetic  
US-09-749-959-49

Query Match 44.2%; Score 50; DB 4; Length 19;  
Best Local Similarity 58.8%; Pred. No. 0.5;  
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 GYPIDDDDYASAGSG 21  
DB 3 GGYPTDVPDYASLGGG 19

## RESULT 6

US-09-285-912A-136  
;/ Sequence 136, Application US/09285912A  
;/ Patent No. 6709814  
;/ GENERAL INFORMATION:  
;/ APPLICANT: Anderson, David  
;/ TITLE OF INVENTION: PEPTIDES CAUSING FORMATION OF COMPACT STRUCTURES  
;/ FILE REFERENCE: A-66103-1/DJB/RMS/SJR  
;/ CURRENT APPLICATION NUMBER: US/09/285,912A  
;/ CURRENT FILING DATE: 1999-04-02  
;/ PRIOR APPLICATION NUMBER: 60/080,444  
;/ PRIOR FILING DATE: 1998-04-02  
;/ NUMBER OF SEQ ID NOS: 150  
;/ SOFTWARE: PatentIn Ver. 2.0  
;/ SEQ ID NO 136  
;/ LENGTH: 18  
;/ TYPE: PRT  
;/ ORGANISM: Artificial sequence  
;/ FEATURE:  
;/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-285-912A-136

Query Match 38.9%; Score 44; DB 4; Length 18;  
Best Local Similarity 60.0%; Pred. No. 3.6;  
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5 GYPIDDDDYASAGSG 19  
DB 3 GGYPTDVPDYASLGG 17

## RESULT 7

US-09-740-876-1  
;/ Sequence 1, Application US/09740876  
;/ Patent No. RE37933  
;/ GENERAL INFORMATION:  
;/ APPLICANT: BRANELLEC, Didier  
;/ WALSH, Kenneth  
;/ ISNER, Jeffrey M.  
;/ DENEFFLE, Patrice  
;/ TITLE OF INVENTION: VIRAL VECTORS AND THEIR USE FOR TREATING  
;/ HYPERPROLIFERATIVE DISORDERS, IN PARTICULAR RESTENOSIS  
;/ NUMBER OF SEQUENCES: 3  
;/ CORRESPONDENCE ADDRESS:  
;/ ADDRESSEE: Rhone-Poulenc Rorer Inc.  
;/ STREET: Mail Drop 3C43, P.O. Box 5093  
;/ CITY: Collegeville  
;/ STATE: PA  
;/ COUNTRY: USA  
;/ ZIP: 19426-0997  
;/ COMPUTER READABLE FORM:  
;/ MEDIUM TYPE: Floppy disk  
;/ OPERATING SYSTEM: PC-DOS/MS-DOS  
;/ SOFTWARE: PatentIn Release #1.0, Version #1.30  
;/ CURRENT APPLICATION NUMBER: US/08/723,726  
;/ FILING DATE: 30-SEP-1996  
;/ CLASSIFICATION: 435  
;/ PRIOR APPLICATION DATA:  
;/ APPLICATION NUMBER: WO PCT/US96/04493  
;/ FILING DATE: 28-MAR-1996  
;/ PRIOR APPLICATION DATA:  
;/ APPLICATION NUMBER: FR 95-04234

;/ MEDIUM TYPE: Floppy disk  
;/ COMPUTER: IBM PC compatible  
;/ OPERATING SYSTEM: PC-DOS/MS-DOS  
;/ SOFTWARE: PatentIn Release #1.0, Version #1.30  
;/ CURRENT APPLICATION DATA:  
;/ APPLICATION NUMBER: US/09/740,876  
;/ FILING DATE: 21-Dec-2000  
;/ CLASSIFICATION: <Unknown>  
;/ PRIOR APPLICATION DATA:  
;/ APPLICATION NUMBER: 08/723,726  
;/ FILING DATE: <Unknown>  
;/ APPLICATION NUMBER: FR 95-04234  
;/ FILING DATE: 31-MAR-1995  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Savitzky, Martin F.  
;/ REGISTRATION NUMBER: 29,699  
;/ REFERENCE/DOCKET NUMBER: ST95022A-US  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: (610)454-3816  
;/ TELEFAX: (610)454-3808  
;/ INFORMATION FOR SEQ ID NO: 1:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 14 amino acids  
;/ TYPE: amino acid  
;/ STRANDEDNESS: <Unknown>  
;/ TOPOLOGY: linear  
;/ MOLECULE TYPE: peptide  
;/ FRAGMENT TYPE: internal  
;/ SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-740-876-1

Query Match 36.3%; Score 41; DB 1; Length 14;  
Best Local Similarity 61.5%; Pred. No. 7.6;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 YPIDDDDYASAGSG 19  
DB 1 YPYDVPDYASLGG 13

## RESULT 8

US-08-723-726-1  
;/ Sequence 1, Application US/08723726  
;/ Patent No. 5851521  
;/ GENERAL INFORMATION:  
;/ APPLICANT: BRANELLEC, Didier  
;/ APPLICANT: WALSH, Kenneth  
;/ APPLICANT: ISNER, Jeffrey M.  
;/ APPLICANT: DENEFFLE, Patrice  
;/ TITLE OF INVENTION: VIRAL VECTORS AND THEIR USE FOR TREATING  
;/ HYPERPROLIFERATIVE DISORDERS, IN PARTICULAR RESTENOSIS  
;/ NUMBER OF SEQUENCES: 3  
;/ CORRESPONDENCE ADDRESS:  
;/ ADDRESSEE: Rhone-Poulenc Rorer Inc.  
;/ STREET: Mail Drop 3C43, P.O. Box 5093  
;/ CITY: Collegeville  
;/ STATE: PA  
;/ COUNTRY: USA  
;/ ZIP: 19426-0997  
;/ COMPUTER READABLE FORM:  
;/ MEDIUM TYPE: Floppy disk  
;/ OPERATING SYSTEM: PC-DOS/MS-DOS  
;/ SOFTWARE: PatentIn Release #1.0, Version #1.30  
;/ CURRENT APPLICATION DATA:  
;/ APPLICATION NUMBER: US/08/723,726  
;/ FILING DATE: 30-SEP-1996  
;/ CLASSIFICATION: 435  
;/ PRIOR APPLICATION DATA:  
;/ APPLICATION NUMBER: WO PCT/US96/04493  
;/ FILING DATE: 28-MAR-1996  
;/ PRIOR APPLICATION DATA:  
;/ APPLICATION NUMBER: FR 95-04234

;; FILING DATE: 31-MAR-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Savitzky, Martin F.  
;; REGISTRATION NUMBER: 29,699  
;; REFERENCE/DOCKET NUMBER: ST95022A-US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (610)454-3816  
;; TELEFAX: (610)454-3808  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 14 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FRAGMENT TYPE: internal  
;;  
US-08-723-726-1

Query Match 36.3%; Score 41; DB 2; Length 14;  
Best Local Similarity 61.5%; Pred. No. 7.6;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 YPIDDDDYASAG 19  
Db 1 YPYDVPDYASLGG 13

RESULT 9  
US-08-390-874C-13  
; Sequence 13, Application US/08390874C  
; Patent No. 6043062  
; GENERAL INFORMATION:  
; APPLICANT: Klippel, Anke  
; APPLICANT: Williams, Lewis T.  
; TITLE OF INVENTION: A Constitutively Active  
; TITLE OF INVENTION: Phosphatidylinositol 3-Kinase and Uses Thereof  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/390,874C  
; FILING DATE: 17-FEB-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dow, Karen B.  
; REGISTRATION NUMBER: 29,684  
; REFERENCE/DOCKET NUMBER: 02307K-057000US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-390-874C-13

Query Match 36.3%; Score 41; DB 3; Length 16;  
Best Local Similarity 61.5%; Pred. No. 8.8;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 YPIDDDDYASAG 19  
Db 1 YPYDVPDYASLGG 13

QY 7 YPIDDDDYASAG 19  
Db 2 YPYDVPDYASLGG 14

RESULT 10  
US-09-265-772-13  
; Sequence 13, Application US/09265772  
; Patent No. 6300111  
; GENERAL INFORMATION:  
; APPLICANT: Klippel, Anke  
; APPLICANT: Williams, Lewis T.  
; TITLE OF INVENTION: A Constitutively Active  
; TITLE OF INVENTION: Phosphatidylinositol 3-Kinase and Uses Thereof  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/265,772  
; FILING DATE: 10-MAR-1999  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/390,874  
; FILING DATE: 17-FEB-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dow, Karen B.  
; REGISTRATION NUMBER: 29,684  
; REFERENCE/DOCKET NUMBER: 02307K-057010US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-265-772-13

Query Match 36.3%; Score 41; DB 3; Length 16;  
Best Local Similarity 61.5%; Pred. No. 8.8;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 YPIDDDDYASAG 19  
Db 2 YPYDVPDYASLGG 14

RESULT 11  
US-09-472-112-4  
; Sequence 4, Application US/09472112  
; Patent No. 6512102  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Lichen  
; APPLICANT: Harrison, Stephen D.  
; APPLICANT: Williams, Lewis T.  
; TITLE OF INVENTION: Compositions and Methods of Diagnosis  
; TITLE OF INVENTION: and Treatment using Casein Kinase I  
; FILE REFERENCE: 35784/193637  
; CURRENT APPLICATION NUMBER: US/09/472,112  
; CURRENT FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: US 60/113,664

; PRIOR FILING DATE: 1998-12-31  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: a C-terminal extension for purification process  
US-09-472-112-4

Query Match 36.3%; Score 41; DB 4; Length 16;  
Best Local Similarity 61.5%; Pred. No. 8.8;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 YPIDDDDDYASAG 19  
||| |||||  
Db 2 YPYDVPDYASLGG 14

RESULT 12  
US-09-500-700-28  
; Sequence 28, Application US/09500700  
; Patent No. 6790941  
; GENERAL INFORMATION:  
; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE  
; APPLICANT: BARBAS III, Carlos F.  
; APPLICANT: GOTTESFELD, Joel M.  
; APPLICANT: WRIGHT, Peter E.  
; TITLE OF INVENTION: ZINC FINGER PROTEIN DERIVATIVES AND METHODS THEREFOR  
; FILE REFERENCE: SCRIPT1160-4  
; CURRENT APPLICATION NUMBER: US/09/500,700  
; CURRENT FILING DATE: 2003-01-10  
; PRIOR APPLICATION NUMBER: US 08/863,813  
; PRIOR FILING DATE: 1997-05-27  
; PRIOR APPLICATION NUMBER: US 08/676,318  
; PRIOR FILING DATE: 1996-12-30  
; PRIOR APPLICATION NUMBER: PCT/US95/00829  
; PRIOR FILING DATE: 1995-01-18  
; PRIOR APPLICATION NUMBER: US 08/312,604  
; PRIOR FILING DATE: 1994-09-28  
; PRIOR APPLICATION NUMBER: US 08/183,119  
; PRIOR FILING DATE: 1994-01-18  
; NUMBER OF SEQ ID NOS: 127  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 28  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-500-700-28

Query Match 36.3%; Score 41; DB 4; Length 25;  
Best Local Similarity 61.5%; Pred. No. 15;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 SGVYPIDDDDDYAS 16  
: ||| |||||  
Db 13 AGQYPYDVPDYAS 25

RESULT 13  
US-09-481-620A-34  
; Sequence 34, Application US/09481620A  
; Patent No. 6506379  
; GENERAL INFORMATION:  
; APPLICANT: ARIAD Gene Therapeutics, Inc.  
; TITLE OF INVENTION: Rapamycin Based Regulation of Biological Events  
; FILE REFERENCE: 345B PCT  
; CURRENT APPLICATION NUMBER: US/09/481,620A  
; CURRENT FILING DATE: 2000-01-12  
; NUMBER OF SEQ ID NOS: 132  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 34

; LENGTH: 26  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: 12CA5epitope--SV40NLS  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(26)  
US-09-481-620A-34

Query Match 36.3%; Score 41; DB 4; Length 26;  
Best Local Similarity 61.5%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 YPIDDDDDYASAG 19  
||| |||||  
Db 5 YPYDVPDYASLGG 17

RESULT 14  
US-08-672-213-39  
; Sequence 39, Application US/08672213  
; Patent No. 6306649  
; GENERAL INFORMATION:  
; APPLICANT: GILMAN, Michael Z.  
; APPLICANT: NATESAN, Sridaran  
; TITLE OF INVENTION: USE OF HETEROLOGOUS TRANSCRIPTION  
; TITLE OF INVENTION: FACTORS IN GENE THERAPY  
; NUMBER OF SEQUENCES: 72  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: ARIAD Gene Therapeutics, Inc.  
; STREET: 26 Landsdowne Street  
; CITY: Cambridge  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02139-4234  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/672,213  
; FILING DATE: 27-JUN-1996  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/000,553  
; FILING DATE: 27-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/019,614  
; FILING DATE: 29-DEC-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BERSTEIN, David L.  
; REGISTRATION NUMBER: 31,235  
; REFERENCE/DOCKET NUMBER: ARIAD 346B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-494-0400  
; TELEFAX: 617-494-0208  
; INFORMATION FOR SEQ ID NO: 39:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 27 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-672-213-39

Query Match 36.3%; Score 41; DB 3; Length 27;  
Best Local Similarity 61.5%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 YPIDDDDDYASAG 19  
||| |||||  
Db 5 YPYDVPDYASLGG 17

RESULT 15  
US-09-781-804-19  
; Sequence 19, Application US/09781804  
; Patent No. 6649595  
; GENERAL INFORMATION:  
; APPLICANT: ARIAD Gene Therapeutics, Inc.  
; TITLE OF INVENTION: Regulation of Biological Events Using No. 6649595el Compounds  
; FILE REFERENCE: 374 USD1  
; CURRENT APPLICATION NUMBER: US/09/781,804  
; CURRENT FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 19  
; LENGTH: 27  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pSMTN3 construct  
; NAME/KEY: DOMAIN  
; LOCATION: (1)-(27)  
; OTHER INFORMATION: pSMTN3+12CA5+SV40T NLS  
US-09-781-804-19

Query Match 36.3%; Score 41; DB 4; Length 27;  
Best Local Similarity 61.5%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 YPIDDDDYASAG 19  
||| |||||  
DB 5 YPYDVPDYASLGG 17

RESULT 16  
US-08-632-514C-29  
; Sequence 29, Application US/08632514C  
; Patent No. 5834234  
; GENERAL INFORMATION:  
; APPLICANT: GALLO, Gregory J.  
; TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BBK  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hale and Dorr  
; STREET: 1455 Pennsylvania Avenue, N.W.  
; CITY: Washington, D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/632,514C  
; FILING DATE: 29-MAY-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WIXON, Henry N.  
; REGISTRATION NUMBER: 32,073  
; REFERENCE/DOCKET NUMBER: 104322.188  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)942-8459  
; TELEFAX: (202)942-8484  
; INFORMATION FOR SEQ ID NO: 29:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-632-514C-29

Query Match 34.5%; Score 39; DB 2; Length 14;

Best Local Similarity 66.7%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 YPIDDDDYASAS 18  
||| |||||  
DB 3 YPYDVPDYASLS 14

RESULT 17  
US-09-188-177-29  
; Sequence 29, Application US/09188177  
; Patent No. 6057132  
; GENERAL INFORMATION:  
; APPLICANT: GALLO, Gregory J.  
; TITLE OF INVENTION: APOPTOSIS ASSOCIATED PROTEIN BBK  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hale and Dorr  
; STREET: 1455 Pennsylvania Avenue, N.W.  
; CITY: Washington, D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/188,177  
; FILING DATE: 2  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WIXON, Henry N.  
; REGISTRATION NUMBER: 32,073  
; REFERENCE/DOCKET NUMBER: 104322.188  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)942-8459  
; TELEFAX: (202)942-8484  
; INFORMATION FOR SEQ ID NO: 29:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 14 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-188-177-29

Query Match 34.5%; Score 39; DB 3; Length 14;  
Best Local Similarity 66.7%; Pred. No. 15;  
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 YPIDDDDYASAS 18  
||| |||||  
DB 3 YPYDVPDYASLS 14

RESULT 18  
US-09-856-920-1  
; Sequence 1, Application US/09856920  
; Patent No. 6740325  
; GENERAL INFORMATION:  
; APPLICANT: Veda Research and Development Co.  
; TITLE OF INVENTION: Peptide-based vaccine for influenza  
; FILE REFERENCE: 9822 PCT  
; CURRENT APPLICATION NUMBER: US/09/856,920  
; CURRENT FILING DATE: 2002-02-19  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 18  
; TYPE: PRT  
; ORGANISM: Influenza virus  
US-09-856-920-1

Query Match 34.5%; Score 39; DB 4; Length 18;  
Best Local Similarity 61.5%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 SGVYPIDDDDYAS 16  
||| |  
Db 5 SNCYPYDVPDYAS 17  
||| |

RESULT 19  
US-09-208-966-15  
; Sequence 15, Application US/09208966  
; Patent No. 6221355  
; GENERAL INFORMATION:  
; APPLICANT: Dowdy, Steven F.  
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF  
; FILE REFERENCE: 48881/1742  
; CURRENT APPLICATION NUMBER: US/09/208,966  
; CURRENT FILING DATE: 1998-12-10  
; EARLIER APPLICATION NUMBER: 60/082,402  
; EARLIER FILING DATE: 1998-04-20  
; EARLIER APPLICATION NUMBER: 60/069,012  
; EARLIER FILING DATE: 1997-12-10  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15  
; TYPE: PRT  
; ORGANISM: human  
US-09-208-966-15

Query Match 33.6%; Score 38; DB 3; Length 16;  
Best Local Similarity 53.8%; Pred. No. 25;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 7 YPIDDDDDYASAG 19  
||| |  
Db 4 YPYDVPDYAGSMG 16  
||| |

RESULT 20  
US-09-775-052A-15  
; Sequence 15, Application US/09775052A  
; Patent No. 6645501  
; GENERAL INFORMATION:  
; APPLICANT: Dowdy, Steven F.  
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF  
; FILE REFERENCE: 48881/1742  
; CURRENT APPLICATION NUMBER: US/09/775,052A  
; CURRENT FILING DATE: 2001-12-05  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/208,966  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-10  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/069,012  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-10  
; NUMBER OF SEQ ID NOS: 57  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: human  
US-09-775-052A-15

Query Match 33.6%; Score 38; DB 4; Length 16;  
Best Local Similarity 53.8%; Pred. No. 25;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 7 YPIDDDDDYASAG 19  
||| |  
Db 4 YPYDVPDYAGSMG 16  
||| |

RESULT 21

US-08-194-560-3  
; Sequence 3, Application US/08194560  
; Patent No. 6255062  
; GENERAL INFORMATION:  
; APPLICANT: Campbell, Judith L.  
; APPLICANT: Budd, Martin E. DNA Polymerases  
; TITLE OF INVENTION: B-Type DNA Polymerases  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
; STREET: 4 Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/194,560  
; FILING DATE: 14-FEB-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Trecartin, Richard F.  
; REGISTRATION NUMBER: 31,801  
; REFERENCE/DOCKET NUMBER: A-59515/RFT/RMS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249  
; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-194-560-3

Query Match 32.7%; Score 37; DB 3; Length 12;  
Best Local Similarity 63.6%; Pred. No. 25;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 VYPIDDDDYAS 16  
:|:|:|  
Db 1 MYPYDVPDYAS 11  
:|:|:|

RESULT 22  
US-08-895-495-40  
; Sequence 40, Application US/08895495  
; Patent No. 6114117  
; GENERAL INFORMATION:  
; APPLICANT: Hepp, Jozsef  
; APPLICANT: Lengyel, Zsolt  
; APPLICANT: Pandey, Rajiv  
; TITLE OF INVENTION: HOMOGENEOUS DIAGNOSTIC ASSAY  
; METHOD UTILIZING SIMULTANEOUS TARGET AND SIGNAL AMPLIFICATION  
; TITLE OF INVENTION: IPLICATION  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 PAGE MILL ROAD  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/895,495  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Axford, Laurie A.  
REGISTRATION NUMBER: 35,053  
REFERENCE/DOCKET NUMBER: 32260-20002.20  
TELEPHONE: 415-813-5600  
TELEFAX: 415-494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 13 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-895-495-40

Query Match 32.7%; Score 37; DB 3; Length 13;  
Best Local Similarity 71.4%; Pred. No. 27;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 YPIDDDDD 13  
DB 6 FPDVDDDD 12

RESULT 23  
US-09-467-997-9  
Sequence 9, Application US/09467997  
Patent No. 6379925  
GENERAL INFORMATION:  
APPLICANT: Kitajewski, Jan  
TITLE OF INVENTION: ANGIOGENIC MODULATION BY NOTCH SIGNAL TRANSDUCTION  
FILE REFERENCE: 53863-A-PCT-US  
CURRENT APPLICATION NUMBER: US/09/467,997  
CURRENT FILING DATE: 1999-12-20  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 18  
TYPE: PRT  
ORGANISM: mouse  
US-09-467-997-9

Query Match 32.7%; Score 37; DB 3; Length 18;  
Best Local Similarity 61.5%; Pred. No. 40;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 SGVYPIDDDDYAS 16  
DB 1 SMAYPYDVPDYAS 13

RESULT 24  
US-08-507-124-18  
Sequence 18, Application US/08507124  
Patent No. 5670618  
GENERAL INFORMATION:  
APPLICANT: McKenzie, Maureen A.  
TITLE OF INVENTION: Insulin-Like Peptide  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolasch & Birch  
STREET: P.O. Box 747  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA

ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/507,124  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/956,342  
FILING DATE: 05-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerald M.  
REGISTRATION NUMBER: 28,977  
REFERENCE/DOCKET NUMBER: 1828-102P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE: Saccharomyces cerevisiae  
US-08-507-124-18

Query Match 32.7%; Score 37; DB 1; Length 23;  
Best Local Similarity 41.2%; Pred. No. 53;  
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 4 SGVYPIDDDDYASASGS 20  
DB 1 AGWPTDCSHYAASAA 17

RESULT 25  
US-08-507-124-33  
Sequence 33, Application US/08507124  
Patent No. 5670618  
GENERAL INFORMATION:  
APPLICANT: McKenzie, Maureen A.  
TITLE OF INVENTION: Insulin-Like Peptide  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Birch, Stewart, Kolasch & Birch  
STREET: P.O. Box 747  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/507,124  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/956,342  
FILING DATE: 05-OCT-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerald M.  
REGISTRATION NUMBER: 28,977



REFERENCE/DOCKET NUMBER: 1828-102P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 23 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Saccharomyces cerevisiae  
US-08-507-124-33

Query Match 32.7%; Score 37; DB 1; Length 23;  
Best Local Similarity 41.2%; Pred. No. 53;  
Matches 7; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 4 SGVYPIDDDDYASAGS 20  
Db 1 AGVWPTDCSHYAAESAA 17

Search completed: April 1, 2005, 21:51:51  
Job time : 44 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2005, 21:50:27 ; Search time 138 Seconds  
(without alignments)  
52.863 Million cell updates/sec

Title: US-10-776-989-9\_COPY\_20\_41  
Perfect score: 113  
Sequence: 1 EASGVYPIDDDYASGSGA 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1413372 seqs, 331592847 residues

Total number of hits satisfying chosen parameters: 336297

Minimum DB seq length: 0  
Maximum DB seq length: 30

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : Published Applications\_AA.\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	44.2	19	9	US-09-749-959-49
2	50	44.2	19	14	US-10-177-725-172
3	50	44.2	19	14	US-10-133-973-107
4	50	44.2	19	15	US-10-393-449-172
5	43	38.1	22	14	US-10-133-973-105
6	42	37.2	15	17	US-10-856-118-34
7	42	37.2	19	17	US-10-851-637-106
8	41	36.3	18	17	US-10-851-637-107
9	41	36.3	19	14	US-10-202-436A-13
10	41	36.3	22	14	US-10-193-934-2
11	41	36.3	25	10	US-09-500-700-28
12	41	36.3	26	15	US-10-341-967-34
13	41	36.3	27	9	US-09-781-804-19
14	36.3				Sequence 19, Appl
15					Sequence 49, Appl
16					Sequence 172, App
17					Sequence 107, App
18					Sequence 172, App
19					Sequence 105, App
20					Sequence 34, Appl
21					Sequence 106, App
22					Sequence 13, Appl
23					Sequence 28, Appl
24					Sequence 34, Appl
25					Sequence 19, Appl
26					Sequence 4, Appl
27					Sequence 51, Appl
28					Sequence 51, Appl
29					Sequence 7, Appl
30					Sequence 27, Appl
31					Sequence 30, Appl
32					Sequence 33, Appl
33					Sequence 34, Appl
34					Sequence 2, Appl
35					Sequence 5, Appl
36					Sequence 8, Appl
37					Sequence 9, Appl
38					Sequence 11, Appl
39					Sequence 12, Appl
40					Sequence 15, Appl
41					Sequence 17, Appl
42					Sequence 2, Appl
43					Sequence 5, Appl
44					Sequence 8, Appl
45					Sequence 9, Appl
46					Sequence 11, Appl
47					Sequence 12, Appl
48					Sequence 15, Appl
49					Sequence 17, Appl
50					Sequence 2, Appl
51					Sequence 5, Appl
52					Sequence 8, Appl
53					Sequence 9, Appl
54					Sequence 11, Appl
55					Sequence 12, Appl
56					Sequence 15, Appl
57					Sequence 17, Appl
58					Sequence 2, Appl
59					Sequence 5, Appl
60					Sequence 8, Appl
61					Sequence 9, Appl
62					Sequence 11, Appl
63					Sequence 12, Appl
64					Sequence 15, Appl
65					Sequence 17, Appl
66					Sequence 2, Appl
67					Sequence 5, Appl
68					Sequence 8, Appl
69					Sequence 9, Appl
70					Sequence 11, Appl
71					Sequence 12, Appl
72					Sequence 15, Appl
73					Sequence 17, Appl
74					Sequence 2, Appl
75					Sequence 5, Appl
76					Sequence 8, Appl
77					Sequence 9, Appl
78					Sequence 11, Appl
79					Sequence 12, Appl
80					Sequence 15, Appl
81					Sequence 17, Appl
82					Sequence 2, Appl
83					Sequence 5, Appl
84					Sequence 8, Appl
85					Sequence 9, Appl
86					Sequence 11, Appl
87					Sequence 12, Appl
88					Sequence 15, Appl
89					Sequence 17, Appl
90					Sequence 2, Appl
91					Sequence 5, Appl
92					Sequence 8, Appl
93					Sequence 9, Appl
94					Sequence 11, Appl
95					Sequence 12, Appl
96					Sequence 15, Appl
97					Sequence 17, Appl
98					Sequence 2, Appl
99					Sequence 5, Appl
100					Sequence 8, Appl

14	41	36.3	27	14	US-10-002-244-33	Sequence 33, Appl
15	41	36.3	27	15	US-10-716-062-19	Sequence 19, Appl
16	39	34.5	18	16	US-10-846-548A-1	Sequence 1, Appl
17	39	34.5	30	15	US-10-375-693-14	Sequence 14, Appl
18	38	33.6	16	10	US-09-134-793-7	Sequence 7, Appl
19	38	33.6	16	10	US-09-775-052-15	Sequence 15, Appl
20	38	33.6	21	16	US-10-772-089-5	Sequence 5, Appl
21	38	33.6	22	9	US-09-864-761-36493	Sequence 36493, A
22	37.5	33.2	17	17	US-10-866-237-18	Sequence 18, Appl
23	37	32.7	20	14	US-10-192-832-44	Sequence 44, Appl
24	37	32.7	30	14	US-10-192-832-58	Sequence 58, Appl
25	37	32.7	30	14	US-10-192-832-59	Sequence 59, Appl
26	36	31.9	10	8	US-08-790-540A-23	Sequence 23, Appl
27	36	31.9	10	8	US-08-791-391A-28	Sequence 28, Appl
28	36	31.9	10	8	US-08-987-689A-7	Sequence 7, Appl
29	36	31.9	10	8	US-08-987-689A-19	Sequence 19, Appl
30	36	31.9	10	8	US-08-987-689A-22	Sequence 22, Appl
31	36	31.9	10	9	US-09-012-135A-21	Sequence 21, Appl
32	36	31.9	10	9	US-09-822-295-13	Sequence 13, Appl
33	36	31.9	10	10	US-09-900-590-28	Sequence 28, Appl
34	36	31.9	10	10	US-09-069-228-13	Sequence 13, Appl
35	36	31.9	10	10	US-09-453-234-30	Sequence 30, Appl
36	36	31.9	10	10	US-09-095-478-20	Sequence 20, Appl
37	36	31.9	10	10	US-09-734-221A-12	Sequence 12, Appl
38	36	31.9	10	11	US-09-784-332-21	Sequence 21, Appl
39	36	31.9	10	14	US-10-193-960-82	Sequence 82, Appl
40	36	31.9	10	14	US-10-292-524-7	Sequence 7, Appl
41	36	31.9	10	14	US-10-292-524-19	Sequence 19, Appl
42	36	31.9	10	14	US-10-292-524-22	Sequence 22, Appl
43	36	31.9	10	14	US-10-305-231-28	Sequence 28, Appl
44	36	31.9	10	14	US-10-016-986-5	Sequence 5, Appl
45	36	31.9	10	15	US-10-452-440-28	Sequence 28, Appl
46	36	31.9	10	15	US-10-464-805-7	Sequence 7, Appl
47	36	31.9	10	15	US-10-464-805-19	Sequence 19, Appl
48	36	31.9	10	15	US-10-464-805-22	Sequence 22, Appl
49	36	31.9	10	15	US-10-463-847-28	Sequence 28, Appl
50	36	31.9	10	15	US-10-284-130-14	Sequence 14, Appl
51	36	31.9	10	17	US-10-849-244-13	Sequence 13, Appl
52	36	31.9	11	14	US-10-128-590-80	Sequence 80, Appl
53	36	31.9	11	14	US-10-177-725-2	Sequence 2, Appl
54	36	31.9	11	14	US-10-128-587A-80	Sequence 80, Appl
55	36	31.9	11	15	US-10-393-449-2	Sequence 2, Appl
56	36	31.9	11	15	US-10-402-364-48	Sequence 48, Appl
57	36	31.9	12	9	US-09-940-063-11	Sequence 11, Appl
58	36	31.9	12	14	US-10-128-590-82	Sequence 82, Appl
59	36	31.9	12	14	US-10-128-587A-82	Sequence 82, Appl
60	36	31.9	12	14	US-10-174-293-11	Sequence 11, Appl
61	36	31.9	15	15	US-10-403-337-10	Sequence 10, Appl
62	36	31.9	15	15	US-10-351-890-10	Sequence 10, Appl
63	36	31.9	16	9	US-09-785-671-4	Sequence 4, Appl
64	36	31.9	17	9	US-09-916-940-51	Sequence 51, Appl
65	36	31.9	17	14	US-10-096-550-51	Sequence 51, Appl
66	36	31.9	17	17	US-10-934-614-51	Sequence 51, Appl
67	36	31.9	18	14	US-10-057-789-7	Sequence 7, Appl
68	36	31.9	18	14	US-10-057-789-27	Sequence 27, Appl
69	36	31.9	18	14	US-10-057-789-30	Sequence 30, Appl
70	36	31.9	18	14	US-10-057-789-33	Sequence 33, Appl
71	36	31.9	18	14	US-10-057-789-34	Sequence 34, Appl
72	36	31.9	18	14	US-10-212-628-7	Sequence 7, Appl
73	36	31.9	18	14	US-10-212-628-27	Sequence 27, Appl
74	36	31.9	18	14	US-10-212-628-30	Sequence 30, Appl
75	36	31.9	18	14	US-10-212-628-33	Sequence 33, Appl
76	36	31.9	18	14	US-10-212-628-34	Sequence 34, Appl
77	36	31.9	18	16	US-10-772-089-2	Sequence 2, Appl
78	36	31.9	19	14	US-10-057-789-5	Sequence 5, Appl
79	36	31.9	19	14	US-10-057-789-8	Sequence 8, Appl
80	36	31.9	19	14	US-10-057-789-9	Sequence 9, Appl
81	36	31.9	19	14	US-10-057-789-10	Sequence 10, Appl
82	36	31.9	19	14	US-10-057-789-11	Sequence 11, Appl
83	36	31.9	19	14	US-10-057-789-12	Sequence 12, Appl
84	36	31.9	19	14	US-10-057-789-15	Sequence 15, Appl
85	36	31.9	19	14	US-10-057-789-17	Sequence 17, Appl
86	36	31.9	19	14	US-10-057-789-19	Sequence 19, Appl

87 36 31.9 19 14 US-10-057-789-20 Sequence 20, Appl  
88 36 31.9 19 14 US-10-057-789-22 Sequence 22, Appl  
89 36 31.9 19 14 US-10-057-789-25 Sequence 25, Appl  
90 36 31.9 19 14 US-10-057-789-28 Sequence 28, Appl  
91 36 31.9 19 14 US-10-057-789-29 Sequence 29, Appl  
92 36 31.9 19 14 US-10-057-789-31 Sequence 31, Appl  
93 36 31.9 19 14 US-10-057-789-32 Sequence 32, Appl  
94 36 31.9 19 14 US-10-057-789-35 Sequence 35, Appl  
95 36 31.9 19 14 US-10-057-789-39 Sequence 39, Appl  
96 36 31.9 19 14 US-10-057-789-40 Sequence 40, Appl  
97 36 31.9 19 14 US-10-212-628-2 Sequence 2, Appl  
98 36 31.9 19 14 US-10-212-628-5 Sequence 5, Appl  
99 36 31.9 19 14 US-10-212-628-8 Sequence 8, Appl  
100 36 31.9 19 14 US-10-212-628-9 Sequence 9, Appl

## ALIGNMENTS

RESULT 1  
US-09-749-959-49  
; Sequence 49, Application US/09749959  
; Publication No. US20010003650A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David  
; APPLICANT: Bogenberger, Jakob M.  
; TITLE OF INVENTION: GREEN FLUORESCENT PROTEIN FUSIONS WITH RANDOM PEPTIDES  
; FILE REFERENCE: A-66900-5/RMS/AMS  
; CURRENT APPLICATION NUMBER: US/09/749,959  
; CURRENT FILING DATE: 2000-12-27  
; PRIOR APPLICATION NUMBER: US 09/169,015  
; PRIOR FILING DATE: 1998-10-08  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 49  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-09-749-959-49

Query Match 44.2%; Score 50; DB 9; Length 19;  
Best Local Similarity 58.8%; Pred. No. 1.5;  
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 GVPIDDDDYASASGSG 21  
DB 3 GGYPTDVPDYASLGGGG 19

RESULT 2  
US-10-177-725-172  
; Sequence 172, Application US/10177725  
; Publication No. US20030143562A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David  
; APPLICANT: Bogenberger, Jakob M.  
; APPLICANT: Peele, Beau R.  
; TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S  
; FILE REFERENCE: A-66900-4/RMS/AMS  
; CURRENT APPLICATION NUMBER: US/10/177,725  
; CURRENT FILING DATE: 2002-06-20  
; PRIOR APPLICATION NUMBER: US 09/415,765  
; PRIOR FILING DATE: 1999-10-08  
; PRIOR APPLICATION NUMBER: US 09/169,015  
; PRIOR FILING DATE: 1998-10-08  
; NUMBER OF SEQ ID NOS: 173  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 172  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial sequence

; FEATURE:  
; OTHER INFORMATION: synthetic  
US-10-177-725-172

Query Match 44.2%; Score 50; DB 14; Length 19;  
Best Local Similarity 58.8%; Pred. No. 1.5;  
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 GVPIDDDDYASASGSG 21  
DB 3 GGYPTDVPDYASLGGGG 19

RESULT 3  
US-10-133-973-107  
; Sequence 107, Application US/10133973  
; Publication No. US20030149254A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS COMPRISING RENILLA GFP  
; FILE REFERENCE: A-66531-3/RMS/CYO  
; CURRENT APPLICATION NUMBER: US/10/133,973  
; CURRENT FILING DATE: 2002-04-24  
; PRIOR APPLICATION NUMBER: US 60/290,287  
; PRIOR FILING DATE: 2001-05-10  
; PRIOR APPLICATION NUMBER: US 09/710,058  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 107  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: HA epitope tag flanked by glycines  
US-10-133-973-107

Query Match 44.2%; Score 50; DB 14; Length 19;  
Best Local Similarity 58.8%; Pred. No. 1.5;  
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 GVPIDDDDYASASGSG 21  
DB 3 GGYPTDVPDYASLGGGG 19

RESULT 4  
US-10-393-449-172  
; Sequence 172, Application US/10393449  
; Publication No. US20030224412A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David  
; APPLICANT: Bogenberger, Jakob M.  
; APPLICANT: Peele, Beau R.  
; TITLE OF INVENTION: STRUCTURALLY BIASED RANDOM PEPTIDE LIBRARIES BASED ON DIFFERENT S  
; FILE REFERENCE: RIGU-007CIP3  
; CURRENT APPLICATION NUMBER: US/10/393,449  
; CURRENT FILING DATE: 2003-03-18  
; PRIOR APPLICATION NUMBER: US 10/177,725  
; PRIOR FILING DATE: 2002-06-20  
; PRIOR APPLICATION NUMBER: US 09/415,765  
; PRIOR FILING DATE: 1999-10-08  
; PRIOR APPLICATION NUMBER: US 09/169,015  
; PRIOR FILING DATE: 1998-10-08  
; NUMBER OF SEQ ID NOS: 173  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 172  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-10-393-449-172

Query Match 44.2%; Score 50; DB 15; Length 19;  
Best Local Similarity 58.8%; Pred. No. 1.5;  
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 GVPYIDDDYASASGSG 21  
DB 3 GGPYDVPDYASLGGGG 19

RESULT 5  
US-10-133-973-105  
; Sequence 105, Application US/10133973  
; Publication No. US20030149254A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS COMPRISING RENILLA GFP  
; FILE REFERENCE: A-68531-3/RMS/CYO  
; CURRENT APPLICATION NUMBER: US/10/133,973  
; CURRENT FILING DATE: 2002-04-24  
; PRIOR APPLICATION NUMBER: US 60/290,287  
; PRIOR FILING DATE: 2001-05-10  
; PRIOR APPLICATION NUMBER: US 09/710,058  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 105  
; LENGTH: 22  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: influenza hemagglutinin epitope tag flanked by linker sequences  
US-10-133-973-105

Query Match 38.1%; Score 43; DB 14; Length 22;  
Best Local Similarity 52.9%; Pred. No. 20;  
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 GVPYIDDDYASASGSG 21  
DB 4 GGPYDVPDYASLGGGG 20

RESULT 6  
US-10-856-118-34  
; Sequence 34, Application US/10856118  
; Publication No. US20050025747A1  
; GENERAL INFORMATION:  
; APPLICANT: Laidlaw, Stephen  
; APPLICANT: Skinner, Mike  
; APPLICANT: Hill, Adrian V.S.  
; APPLICANT: Gilbert, Sarah C.  
; APPLICANT: Anderson, Richard  
; TITLE OF INVENTION: Vaccine  
; FILE REFERENCE: 3742.1000-000  
; CURRENT APPLICATION NUMBER: US/10/856,118  
; CURRENT FILING DATE: 2004-05-27  
; PRIOR APPLICATION NUMBER: PCT/GB02/005411  
; PRIOR FILING DATE: 2002-12-02  
; PRIOR APPLICATION NUMBER: GB0128733.3  
; PRIOR FILING DATE: 2001-11-30  
; PRIOR APPLICATION NUMBER: US 60/334,649  
; PRIOR FILING DATE: 2001-11-30  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 34  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Flanking sequence  
US-10-856-118-34

Query Match 37.2%; Score 42; DB 17; Length 15;  
Best Local Similarity 57.1%; Pred. No. 18;  
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 SGVYPIDDDDYASA 17  
DB 2 TNVYDVPDYAAA 15

RESULT 7  
US-10-851-637-106  
; Sequence 106, Application US/10851637  
; Publication No. US20050015039A1  
; GENERAL INFORMATION:  
; APPLICANT: Salzwedel, Karl  
; APPLICANT: Li, Feng  
; APPLICANT: Wild, Carl T.  
; APPLICANT: Allaway, Graham P.  
; APPLICANT: Freed, Eric O.  
; TITLE OF INVENTION: Inhibition of HIV-1 Replication by Distribution of the Processing  
; TITLE OF INVENTION: of the Viral Capsid-Spacer Peptide 1 Protein  
; FILE REFERENCE: 1900.0430003  
; CURRENT APPLICATION NUMBER: US/10/851,637  
; CURRENT FILING DATE: 2004-05-24  
; PRIOR APPLICATION NUMBER: US 10/766,528  
; PRIOR FILING DATE: 2004-01-29  
; PRIOR APPLICATION NUMBER: US 60/496,660  
; PRIOR FILING DATE: 2003-08-21  
; PRIOR APPLICATION NUMBER: US 60/443,180  
; PRIOR FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 115  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 106  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: CA-SPI in HIV-1 NL4-3 with HA epitope  
US-10-851-637-106

Query Match 37.2%; Score 42; DB 17; Length 19;  
Best Local Similarity 56.2%; Pred. No. 24;  
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 BEASGVYPIDDDDYAS 16  
DB 2 EAMSVYDVPDYAT 17

RESULT 8  
US-10-851-637-107  
; Sequence 107, Application US/10851637  
; Publication No. US20050015039A1  
; GENERAL INFORMATION:  
; APPLICANT: Salzwedel, Karl  
; APPLICANT: Li, Feng  
; APPLICANT: Wild, Carl T.  
; APPLICANT: Allaway, Graham P.  
; APPLICANT: Freed, Eric O.  
; TITLE OF INVENTION: Inhibition of HIV-1 Replication by Distribution of the Processing  
; TITLE OF INVENTION: of the Viral Capsid-Spacer Peptide 1 Protein  
; FILE REFERENCE: 1900.0430003  
; CURRENT APPLICATION NUMBER: US/10/851,637  
; CURRENT FILING DATE: 2004-05-24  
; PRIOR APPLICATION NUMBER: US 10/766,528  
; PRIOR FILING DATE: 2004-01-29  
; PRIOR APPLICATION NUMBER: US 60/496,660  
; PRIOR FILING DATE: 2003-08-21  
; PRIOR APPLICATION NUMBER: US 60/443,180  
; PRIOR FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 115  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 107

```
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CA-SP1 in HIV-1 NL4-3 with HA epitope
US-10-851-637-107

Query Match      36.3%; Score 41; DB 17; Length 18;
Best Local Similarity 60.0%; Pred. No. 31;
Matches 9; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 EASGVYPIDDDYA 15
Db 2 EAMQVYPYDVPDYA 16

RESULT 9
US-10-202-436A-13
; Sequence 13, Application US/10202436A
; Publication No. US20030049261A1
; GENERAL INFORMATION:
; APPLICANT: ELIAS, JACK A
; APPLICANT: ZHU, ZHOU
; TITLE OF INVENTION: METHODS COMPOSITIONS AND KITS RELATING TO CHITINASES AND CHITINASE
; FILE REFERENCE: 044574-5107
; CURRENT APPLICATION NUMBER: US/10/202,436A
; PRIOR FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: U.S. 60/307,432
; PRIOR FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 13
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Anti-AMCase Antibody Immunizing Peptide
US-10-202-436A-13

Query Match      36.3%; Score 41; DB 14; Length 19;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EASGVYPIDDD 12
Db 2 DRADGLYPVADD 13

RESULT 10
US-10-193-934-2
; Sequence 2, Application US/10193934
; Publication No. US20030055018A1
; GENERAL INFORMATION:
; APPLICANT: Chan, Tung O
; APPLICANT: Philip, Tschlis N
; TITLE OF INVENTION: A Small GTPase and Tyrosine Kinase Co-regulated Molecular Switch
; FILE REFERENCE: Phosphoinositide 3-Kinase Regulated Subunit
; CURRENT APPLICATION NUMBER: US/10/193,934
; CURRENT FILING DATE: 2002-07-11
; PRIOR APPLICATION NUMBER: US 60/304,498
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: US 60/363,078
; PRIOR FILING DATE: 2002-03-11
; PRIOR APPLICATION NUMBER: US 60/308,654
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: An extended Hemagglutinin epitope
US-10-193-934-2

Query Match      36.3%; Score 41; DB 14; Length 22;
Best Local Similarity 61.5%; Pred. No. 39;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 YPIDDDDYASAG 19
Db 5 YPYDVPDYASLGG 17

RESULT 11
US-09-500-700-28
; Sequence 28, Application US/09500700
; Publication No. US20030059767A1
; GENERAL INFORMATION:
; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE
; APPLICANT: BARBAS III, Carlos F.
; APPLICANT: GOTTESFELD, Joel M.
; APPLICANT: WRIGHT, Peter E.
; TITLE OF INVENTION: ZINC FINGER PROTEIN DERIVATIVES AND METHODS THEREFOR
; FILE REFERENCE: SCRIP1160-4
; CURRENT APPLICATION NUMBER: US/09/500,700
; CURRENT FILING DATE: 2003-01-10
; PRIOR APPLICATION NUMBER: US 08/863,813
; PRIOR FILING DATE: 1997-05-27
; PRIOR APPLICATION NUMBER: US 08/676,318
; PRIOR FILING DATE: 1996-12-30
; PRIOR APPLICATION NUMBER: PCT/US95/00829
; PRIOR FILING DATE: 1995-01-18
; PRIOR APPLICATION NUMBER: US 08/312,604
; PRIOR FILING DATE: 1994-09-28
; PRIOR APPLICATION NUMBER: US 08/183,119
; PRIOR FILING DATE: 1994-01-18
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 28
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Mouse
US-09-500-700-28

Query Match      36.3%; Score 41; DB 10; Length 25;
Best Local Similarity 61.5%; Pred. No. 45;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 SGVYPIDDDYAS 16
Db 13 AGQYPYDVPDYAS 25

RESULT 12
US-10-341-967-34
; Sequence 34, Application US/10341967
; Publication No. US20030206891A1
; GENERAL INFORMATION:
; APPLICANT: ARIAD Gene Therapeutics, Inc.
; TITLE OF INVENTION: Rapamycin Based Regulation of Biological Events
; FILE REFERENCE: 34SB PCT
; CURRENT APPLICATION NUMBER: US/10/341,967
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US/09/481,620A
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 132
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 34
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 12CA5epitope--SV40NLS
```

;  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(26)  
US-10-341-967-34

Query Match 36.3%; Score 41; DB 15; Length 26;  
Best Local Similarity 61.5%; Pred. No. 47;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 YPIDDDDYASAG 19  
||| |||||  
DB 5 YPYDVPDYASLGG 17

## RESULT 13

US-09-781-804-19  
; Sequence 19, Application US/09781804  
; Patent No. US20020107189A1  
; GENERAL INFORMATION:  
; APPLICANT: ARIAD Gene Therapeutics, Inc.  
; TITLE OF INVENTION: Regulation of Biological Events Using No. US20020107189A1e1 Compo  
; FILE REFERENCE: 374 USD1  
; CURRENT APPLICATION NUMBER: US/09/781,804  
; CURRENT FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 19  
; LENGTH: 27  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pSMTN3 construct  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(27)  
; OTHER INFORMATION: pSMTN3+12CA5+SV40T NLS  
US-09-781-804-19

Query Match 36.3%; Score 41; DB 9; Length 27;  
Best Local Similarity 61.5%; Pred. No. 49;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 YPIDDDDYASAG 19  
||| |||||  
DB 5 YPYDVPDYASLGG 17

## RESULT 14

US-10-002-244-33  
; Sequence 33, Application US/10002244  
; Publication No. US20030143731A1  
; GENERAL INFORMATION:  
; APPLICANT: ARIAD Gene Therapeutics, Inc.  
; TITLE OF INVENTION: Use of Heterologous Transcription Factors in Gene Therapy  
; FILE REFERENCE: 346B USCL  
; CURRENT APPLICATION NUMBER: US/10/002,244  
; CURRENT FILING DATE: 2002-01-29  
; NUMBER OF SEQ ID NOS: 68  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 33  
; LENGTH: 27  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(27)  
; OTHER INFORMATION: pSMTN3  
US-10-002-244-33

Query Match 36.3%; Score 41; DB 14; Length 27;  
Best Local Similarity 61.5%; Pred. No. 49;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 YPIDDDDYASAG 19

DB 5 YPYDVPDYASLGG 17  
||| |||||

## RESULT 15

US-10-716-062-19  
; Sequence 19, Application US/10716062  
; Publication No. US20040082515A1  
; GENERAL INFORMATION:  
; APPLICANT: ARIAD Gene Therapeutics, Inc.  
; TITLE OF INVENTION: Regulation of Biological Events Using Novel Compounds  
; FILE REFERENCE: 374 USD1  
; CURRENT APPLICATION NUMBER: US/10/716,062  
; CURRENT FILING DATE: 2003-11-18  
; PRIOR APPLICATION NUMBER: US/09/781,804  
; PRIOR FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 19  
; LENGTH: 27  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pSMTN3 construct  
; NAME/KEY: DOMAIN  
; LOCATION: (1)..(27)  
; OTHER INFORMATION: pSMTN3+12CA5+SV40T NLS  
US-10-716-062-19

Query Match 36.3%; Score 41; DB 15; Length 27;  
Best Local Similarity 61.5%; Pred. No. 49;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 YPIDDDDYASAG 19  
||| |||||  
DB 5 YPYDVPDYASLGG 17

## RESULT 16

US-10-846-548A-1  
; Sequence 1, Application US/10846548A  
; Publication No. US20050002954A1  
; GENERAL INFORMATION:  
; APPLICANT: ARNON, Ruth  
; APPLICANT: BEN-YEDIDIA, Tamar  
; APPLICANT: LEVI, Raphael  
; TITLE OF INVENTION: PEPTIDE-BASED VACCINE FOR INFLUENZA  
; FILE REFERENCE: 27811  
; CURRENT APPLICATION NUMBER: US/10/846,548A  
; CURRENT FILING DATE: 2004-05-17  
; PRIOR FILING DATE: 2004-05-17  
; PRIOR APPLICATION NUMBER: US 09/856,920  
; PRIOR FILING DATE: 2001-05-29  
; PRIOR APPLICATION NUMBER: PCT/IL99/00640  
; PRIOR FILING DATE: 1999-11-28  
; PRIOR APPLICATION NUMBER: IL 127331  
; PRIOR FILING DATE: 1998-11-30  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 18  
; TYPE: PRT  
; ORGANISM: Influenza virus  
US-10-846-548A-1

Query Match 34.5%; Score 39; DB 16; Length 18;  
Best Local Similarity 61.5%; Pred. No. 62;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 SGVPIDDDDYAS 16  
||| |||||  
DB 5 SNCYPYDVPDYAS 17

```
RESULT 17
US-10-375-693-14
; Sequence 14, Application US/10375693
; Publication No. US2004003873A1
; GENERAL INFORMATION:
; APPLICANT: Floman, Harvey
; APPLICANT: Jungnickel, Melissa
; APPLICANT: Sutton, Keith
; TITLE OF INVENTION: ENKURIN AND USES THEREOF
; FILE REFERENCE: 07917-159001
; CURRENT APPLICATION NUMBER: US/10/375,693
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: US 60/359,870
; PRIOR FILING DATE: 2002-02-25
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetically generated peptide
US-10-375-693-14

Query Match      34.5%; Score 39; DB 15; Length 30;
Best Local Similarity 63.6%; Pred. No. 1.le+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      5 GVPYDDDDYA 15
       1:|||||
Db      20 GMPYDVPDYA 30

RESULT 18
US-09-134-793-7
; Sequence 7, Application US/09134793
; Publication No. US20030040038A1
; GENERAL INFORMATION:
; APPLICANT: Dowdy, Steven F.
; APPLICANT: Jesses, Joel A.
; TITLE OF INVENTION: INDUCIBLE REGULATORY SYSTEM
; TITLE OF INVENTION: AND USE THEREOF
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dike, Bronstein, Roberts & Cushman, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,793
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/056,713
; FILING DATE: 22-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Corless, Peter F
; REGISTRATION NUMBER: 33,860
; REFERENCE/DOCKET NUMBER: 47275
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 7:

Query Match      33.6%; Score 38; DB 10; Length 16;
Best Local Similarity 53.8%; Pred. No. 77;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      7 YPIDDDDYASAG 19
       1:|||||
Db      4 YPYDVPDYAGSMG 16

RESULT 19
US-09-775-052-15
; Sequence 15, Application US/09775052
; Publication No. US20030054000A1
; GENERAL INFORMATION:
; APPLICANT: Dowdy, Steven F.
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
; FILE REFERENCE: 48881/1742
; CURRENT APPLICATION NUMBER: US/09/775,052
; CURRENT FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 09/208,966
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: 60/082,402
; PRIOR FILING DATE: 1998-04-20
; PRIOR APPLICATION NUMBER: 60/069,012
; PRIOR FILING DATE: 1997-12-10
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 15
; LENGTH: 16
; TYPE: PRT
; ORGANISM: human
US-09-775-052-15

Query Match      33.6%; Score 38; DB 10; Length 16;
Best Local Similarity 53.8%; Pred. No. 77;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      7 YPIDDDDYASAG 19
       1:|||||
Db      4 YPYDVPDYAGSMG 16

RESULT 20
US-10-772-089-5
; Sequence 5, Application US/10772089
; Publication No. US20040192609A1
; GENERAL INFORMATION:
; APPLICANT: The Brigham & Women's Hospital, Inc.
; APPLICANT: Farzan, Michael R
; APPLICANT: Dorfman, Tatyana K
; TITLE OF INVENTION: Peptides Binding gp120 of HIV-1
; FILE REFERENCE: 7570/80211
; CURRENT APPLICATION NUMBER: US/10/772,089
; CURRENT FILING DATE: 2004-02-05
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-772-089-5

Query Match      33.6%; Score 38; DB 16; Length 21;
Best Local Similarity 58.3%; Pred. No. 1e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```



QY 10 DDDYASASGSG 21  
:|||||  
Db 7 DNDYDRSGRG 18  
:|||||

RESULT 21  
US-09-864-761-36493  
; Sequence 36493, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn; Sharon G.  
; APPLICANT: Rank; David R.  
; APPLICANT: Hanzel; David K.  
; APPLICANT: Chen; Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeonica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 36493  
; LENGTH: 22  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC011235.2  
; OTHER INFORMATION: EXPRESSED IN PETAL LIVER, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.97  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3

; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1  
US-09-864-761-36493

Query Match 33.6%; Score 38; DB 9; Length 22;  
Best Local Similarity 53.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 10 DDDYASASGSGA 22  
:|||||  
Db 9 DDDDDGGGGSGS 21  
:|||||

RESULT 22  
US-10-866-237-18  
; Sequence 18, Application US/10866237  
; Publication No. US20050037392A1  
; GENERAL INFORMATION:  
; APPLICANT: Griffiths; Andrew  
; APPLICANT: Tawfik; Dan  
; APPLICANT: Sepp; Armin  
; TITLE OF INVENTION: OPTICAL SORTING METHOD  
; FILE REFERENCE: 18936/2023 (21465-202)  
; CURRENT APPLICATION NUMBER: US/10/866,237  
; CURRENT FILING DATE: 2004-06-11  
; PRIOR APPLICATION NUMBER: US 09/896,915  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: PCT/GB00/00030  
; PRIOR FILING DATE: 2000-01-06  
; PRIOR APPLICATION NUMBER: GB 9900298.2  
; PRIOR FILING DATE: 1999-01-07  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 18  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Plasmid  
US-10-866-237-18

Query Match 33.2%; Score 37.5; DB 17; Length 17;  
Best Local Similarity 45.0%; Pred. No. 98;  
Matches 9; Conservative 2; Mismatches 6; Indels 3; Gaps 1;

QY 3 ASGVPIDDDDYASASGSGA 22  
:|||||  
Db 1 SCGGYPYDVPDYA---GGGS 17  
:|||||

RESULT 23  
US-10-192-832-44  
; Sequence 44, Application US/10192832  
; Publication No. US20030176335A1  
; GENERAL INFORMATION:  
; APPLICANT: ZHANG; SHUANG  
; APPLICANT: VAUTHEY; SYLVAIN  
; TITLE OF INVENTION: SURFACTANT PEPTIDE NANOSTRUCTURES, AND USES THEREOF  
; FILE REFERENCE: MTV-043.01  
; CURRENT APPLICATION NUMBER: US/10/192,832  
; CURRENT FILING DATE: 2002-07-10  
; PRIOR APPLICATION NUMBER: 60/304,256  
; PRIOR FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 76  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 44  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-192-832-44

Query Match 32.7%; Score 37; DB 14; Length 20;  
Best Local Similarity 53.8%; Pred. No. 1.4e+02;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Search completed: April 1, 2005, 22:03:01  
Job time : 139 secs

QY 10 DDDYASASGSGA 22  
| | | | | : | : |  
Db 7 DDDDDAAAAAAA 19

## RESULT 24

US-10-192-832-58  
; Sequence 58, Application US/10192832  
; Publication No. US20030176335A1  
; GENERAL INFORMATION:  
; APPLICANT: ZHANG, SHUGUANG  
; APPLICANT: VAUTHEY, SYLVAIN  
; TITLE OF INVENTION: SURFACTANT PEPTIDE NANOSTRUCTURES, AND USES THEREOF  
; FILE REFERENCE: MTV-043.01  
; CURRENT APPLICATION NUMBER: US/10/192,832  
; CURRENT FILING DATE: 2002-07-10  
; PRIOR APPLICATION NUMBER: 60/304,256  
; PRIOR FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 76  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 58  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-192-832-58

Query Match 32.7%; Score 37; DB 14; Length 30;  
Best Local Similarity 53.8%; Pred. No. 2.2e+02;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 10 DDDYASASGSGA 22  
| | | | | : | : |  
Db 7 DDDDDAAAAAAA 19

## RESULT 25

US-10-192-832-59  
; Sequence 59, Application US/10192832  
; Publication No. US20030176335A1  
; GENERAL INFORMATION:  
; APPLICANT: ZHANG, SHUGUANG  
; APPLICANT: VAUTHEY, SYLVAIN  
; TITLE OF INVENTION: SURFACTANT PEPTIDE NANOSTRUCTURES, AND USES THEREOF  
; FILE REFERENCE: MTV-043.01  
; CURRENT APPLICATION NUMBER: US/10/192,832  
; CURRENT FILING DATE: 2002-07-10  
; PRIOR APPLICATION NUMBER: 60/304,256  
; PRIOR FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 76  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 59  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-10-192-832-59

Query Match 32.7%; Score 37; DB 14; Length 30;  
Best Local Similarity 53.8%; Pred. No. 2.2e+02;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 10 DDDYASASGSGA 22  
| | | | | : | : |  
Db 17 DDDDDAAAAAAA 29